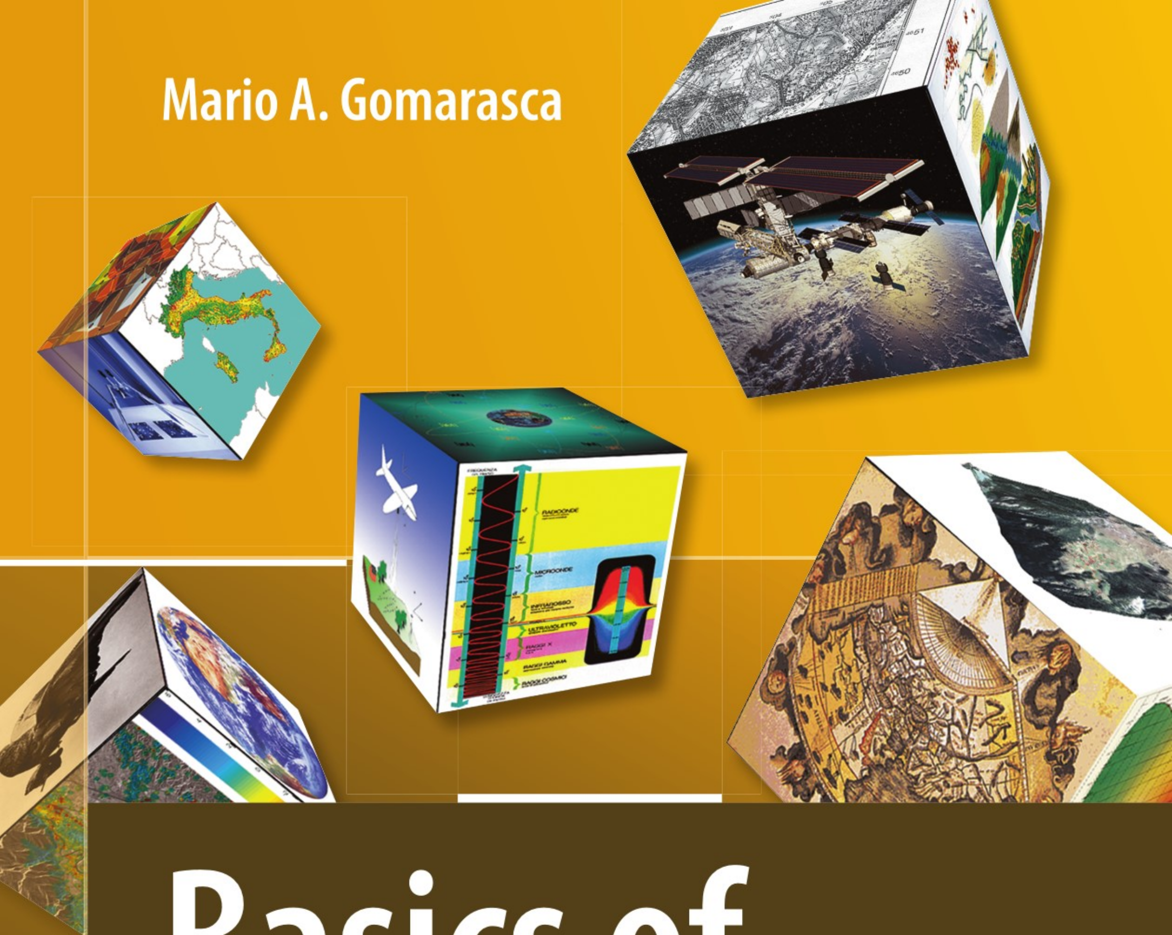
Mario A. Gomarasca



Basics of Geomatics

Springer

Basics of Geomatics

Basics of Geomatics

by

Mario A. Gomarasca

National Research Council of Italy, Institute for the
Electromagnetic Sensing of the Environment,
Milano, Italy

Springer

Mario A. Gomarasca
National Research Council of Italy
Institute for the Electromagnetic
Sensing of the Environment
Via Bassini, 15
20133 Milano
Italy
gomarasca.m@irea.cnr.it

ISBN 978-1-4020-9013-4 e-ISBN 978-1-4020-9014-1
DOI 10.1007/978-1-4020-9014-1
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009926868

This is a translation revised and enlarged of the original work in Italian "Elementi di Geomatica" published by Associazione Italiana di Telerilevamento, © 2004.

Cover illustration: Images by Luca Di Ionno

English translation by the Author assisted by Sara de Santis and Andrew Lowe

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

This project is for my sons Ilaria Camilla (Ila)
and Jacopo Andrea (Jepus)

Foreword

Geomatics is a neologism, the use of which is becoming increasingly widespread, even if it is not still universally accepted. It includes several disciplines and techniques for the study of the Earth's surface and its environments, and computer science plays a decisive role. A more meaningful and appropriate expression is *Geo-spatial Information* or *GeoInformation*.

Geo-spatial Information embeds topography in its more modern forms (measurements with electronic instrumentation, sophisticated techniques of data analysis and network compensation, global satellite positioning techniques, laser scanning, etc.), analytical and digital photogrammetry, satellite and airborne remote sensing, numerical cartography, geographical information systems, decision support systems, WebGIS, etc.

These specialized fields are intimately interrelated in terms of both the basic science and the results pursued: rigid separation does not allow us to discover several common aspects and the fundamental importance assumed in a search for solutions in the complex survey context.

The objective pursued by Mario A. Gomarasca, one that is only apparently modest, is to publish an integrated text on the surveying theme, containing simple and comprehensible concepts relevant to experts in Geo-spatial Information and/or specifically in one of the disciplines that compose it. At the same time, the book is rigorous and synthetic, describing with precision the main instruments and methods connected to the multiple techniques available today.

The book is addressed not to super-specialists, but to a wider group of technicians and students who may use Geo-spatial Information in their work, or who already use it as part of their daily professional activity or study. More specifically the book targets at land managers, operating in natural or anthropic environments (engineers, geologists, agronomists, architects, urban planners, operating in the field of architectural assets and environment, technicians at land-surveying agencies, etc.), and students at both first and master levels, more and more of whom are facing themes in which the disciplines of the survey play a determining role.

Mario A. Gomarasca is a researcher at the National Research Council of Italy, expert in remote sensing applied to agriculture and environment, and more recently, for many years (1997–2003), he has held the prestigious and engaging position of president of ASITA (Federation of the Scientific Associations for Land and

Environment Information). In this role, which he performs with enthusiasm and great efficiency, he has become a privileged observer of the topics of Geo-spatial Information, since he coordinates the National Conferences, at which some hundreds of scientific papers are presented annually.

The absolute specialist in a single field will not find profound or very specific innovative elements in his/her particular competence, but the same specialist will be able to add elements from adjacent, interrelated disciplines and techniques.

The readers, whether university student, professional, technician or lay student, will find ready access to the fundamental concepts and up-to-date information on the state of the art, giving them a wider field of view of the complex, multidisciplinary problems related to land surveying and the environment, especially in land planning.

This objective, which I must warn the reader is decidedly other than modest, is totally achieved in this book. To both, the book and its author Mario A. Gomarasca, I wish the best and all the good fortune they deserve.

Turin, Italy Sergio Dequal

Author's Preface

When I decided to revisit my first book 'GIS and Remote Sensing for the Management of the Agricultural and Environment Resources' (published by AIT in 1997) at the end of November 1999, while waiting for a flight to Niamey, Niger, at the Paris Charles De Gaulle airport, with my unforgettable colleague and friend Eugenio Zilioli, I had in mind to only update the text for the second reprint.

While reading and re-reading, with the growing knowledge that in the meantime was integrating itself in ASITA, the Italian Federation of the Scientific Associations for the Land and Environment Information, of which I had been elected and serving as President since 1998, and with the rising interests of the profession, grew the idea to broaden the content and to develop a more ambitious project involving an interactive approach to some of the main topics of Geo-spatial Information.

After years of reading, research, study, complex bibliographical consultation and selection, along with thorough and sometime critical reviews from many experts, the Italian version of the book (2004) introduced a panorama to the neophyte and completed the framework for those who already work in the field, in order to integrate the knowledge of Geo-spatial Information.

Considering the large success of the book in Italy and the worldwide interest and development of geomatics, I decided to undertake the challenge in preparing the English revised and enlarged version of that book.

This book introduces various disciplines and techniques and is offered as a review of the subject to stimulate the reader's interests. Mathematical demonstrations and deeper explanations have been omitted, but an accurate selected bibliography is provided, chapter by chapter, to assist in finding specific references.

Geo-spatial Information is still a relatively new discipline with fuzzy contours, open to many interpretations; adding my own personal point of view, which could generate approval and criticism, opening, I hope, a scientific and professional constructive debate.

The book does not lay claim to answering multiple issues that Geomatics includes, but it proposes an interdisciplinary integration in order to contribute to face the problems provided by this complex world.

The necessity of defining technical terms occurs in many passages of this book; I tried to impose an order on the labyrinth of definitions and acronyms that are often used in a general way. At this stage, with no existing universally recognized ontological dictionary and thesaurus, I have selected a nomenclature with the more

commonly used definitions and which hopefully mediates between sometimes contrasting positions.

The book is aimed at those who await an introduction and a broadening of the disciplines and techniques of Geomatics (Geo-spatial Information), with particular attention to public administration, university students, training courses and professionals.

Several people have helped me in preparing the first Italian version of this book, in particular, the Italian Remote Sensing Association (AIT), the Institute for the Electromagnetic Sensing of the Environment in Milan (CNR-IREA) and Giovanni Lechi, Polytechnic of Milan have played a fundamental role and I thank all with special affection, as well as the Department of Engineering of the Territory, Environment and Geotechnology of the Polytechnic of Turin, my second professional family.

Substantial support was provided by Rainer Reuter, EARSeL (European Association of the Remote Sensing Laboratories) chairman; Sandro Annoni, JRC, European Commission, Ispra; Antonio Di Gregorio, FAO Africover Plan, Nairobi, Kenya; Jimmy Johnston, National Wetlands Research Center, USGS, Lafayette, USA; Richard Escadafal, CESBIO, Toulouse, F; Ramon Norberto Fernandez, UNDP-GRID, Nairobi, Kenya; Guy Weets, formerly DG Information Society, European Commission, Brussels; Daniele Rizzi, Geographical Information System of the Commission (GISCO), Eurostat, European Commission, Brussels; Luciano Surace, Italian Hydrographical Institute of Navy and ASITA president; Giuseppe Scanu, University of Sassari and Italian Cartographic Association (AIC) president; Ruggero Casacchia, National Research Council of Rome and Italian Remote Sensing Association (AIT) president and Mauro Salvemini, University La Sapienza of Rome, AM/FM Geographical Information System and EUROGI (European Umbrella Organization for Geographic Information) president.

Moreover, I thank Claudio Prati and Fabio Rocca, Department of Electronics and Information, Polytechnic of Milan; Italian Space Agency (ASI); European Space Agency (ESA); Remote Sensing Europe of Milan; Compagnia Generale Ripreseaeree of Parma (CGR) Blom ASA Group and Agronomic Institute for Overseas (IAO), Florence, Italy for assistance with documentation and images.

I thank the several reviewers and advisors, fundamentals with their professionalism and competence. A special acknowledgement goes to Chris J. Johannsen, Professor Emeritus of Agronomy, my tutor during my stay (1988–1989) as visiting scientist at the Laboratory for the Application of Remote Sensing (LARS), Purdue University, West Lafayette, Indiana, USA.

Milan, Italy Mario A. Gomarasca

Contents

List of Colour Plates

Acronyms

Acronym	Definition
A	
A/D	*Analog to Digital*
AATSR	*Advanced Along-Track Scanning Radiometer*
ACM	*Association for Computing Machinery*
ADEO	*Advanced Earth Observing Satellite*
ADEOS	*Advanced Earth Observing System (Japan)*
ADG	*Africover Database Gateway*
ADM	*Atmospheric Dynamics Mission*
ADRG	*ARC-Digitized Raster Graphics*
ADS40	*Airborne Digital Sensor 40*
ADSL	*Asymmetric Digital Subscriber Line*
AEB	*Agência Espacial Brasileira*
AERONET	*AErosol RObotic NETwork (NASA)*
AGI	*Advanced Global Imager (NASA)*
AI	*Artificial Intelligence*
AID	*Africover Interactive Database*
AIMs	*Africover Interpretation and Mapping System*
AIMS	*Airborne Integrated Mapping System*
AIRS	*Atmospheric Infrared Sounder (NASA EOS)*
ALI	*Advanced Land Imager (USA)*
ALOS	*Advanced Land Observing Satellite (Japan)*
ALS	*Aerial (Airborne) Laser Scanning*
ALTM	*Airborne Laser Terrain Mapper*
ALTMS	*Airborne Laser Topographic Mapping System*
ALU	*Arithmetic–Logic Unit*
AM	*Automated Mapping*
AM/FM	*Automated Mapping/Facilities Management*
AMI SAR	*Active Microwave Imager Synthetic Aperture Radar*
AMPS	*Automatic Mapping and Planning System*
AMSD	*Adjusted Mapping Support Data*
AMSR	*Advanced Microwave Scanning Radiometer (Japan; NASA EOS)*
AMSU	*Advanced Microwave Sounding Unit (NASA EOS)*
ANOVA	*Analysis of Variance*

Acronym	Definition
AOIPS	*Atmospheric and Oceanographic Image Processing System*
APCM	*Aerial Photography Contract Management System*
APFO	*Aerial Photography Field Office (USDA)*
API	*Application Program Interface*
APIS	*Aerial Photography Information System*
APMI	*Aerial Photography Micrographic Index System (USGS)*
APQF	*Aerial Photography Quad File*
APR	*Airborne Profile Recorder*
APSR	*Aerial Photography Summary Record*
APSRS	*Aerial Photography Summary Record System (USGS)*
APTS	*Aerial Profiling of Terrain System*
AQT	*Association Québecoise de Télédétection*
ARNS	*Aeronautical Radio Navigation Service*
ARVI	*Atmospherically Resistant Vegetation Index*
AS	*Applicative Software*
ASAR	*Advanced Synthetic Aperture Radar; Aerial Synthetic Aperture Radar*
ASC	*Agence Spatiale Canadienne*
ASCAT	*MetOp's Advanced SCATterometer*
ASCIE	*American Standard Code for Information Exchange*
ASCII	*American Standard Code for Information Interchange*
ASI	*Agenzia Spaziale Italiana Italian Space Agency*
ASP	*Active Server Pages; Application Service Provider*
ASTER	*Advanced Spaceborne Thermal Emission and Reflection Radiometer (Japan; NASA EOS)*
ATBD	*Algorithm Theoretical Basis Document (NASA EOS)*
ATM	*Airborne Terrain Mapper*
ATMOS	*Atmospheric and Ocean Observation Series (Japan)*
ATR1	*Automatic Target Recognition Module (Leica)*
ATSR	*Along-Track Scanning Radiometer (ESA ERS)*
ATV	*Automated Transfer Vehicle*
AVHRR	*Advanced Very High Resolution Radiometer (NOAA)*
AVI	*Audio Visual Interleave*
AVIRIS	*Advanced Visible/InfraRed Imaging Spectrometer*
AVL	*Automatic Vehicle Location*
AVNIR	*Advanced Visible and Near-Infrared Radiometer (Japan, ADEOS)*
AWARE	*A tool for monitoring and forecasting Available WAter REsource in mountain environment*
AWiFS	*Advanced Wide Field Sensor*
B	
BE	*Best Effort*
B/W	*Black and White*
BWA	*Broadband Wireless Access*
BARCIS	*BARCode Information System*
BGS	*British Geological Survey*
BIH	*Bureau International de l'Heure*
BIL	*Band Interleaved by Line*
BIOS	*Basic Input/Output System*
BIP	*Band Interleaved by Pixel*
BIPM	*Bureau International des Poids et Mesures*

Acronym	Definition
BLG	*Binary Line Generalization*
BMP	*Bitmapped Image Format*
BNSC	*British National Space Centre (UK)*
BOREAS	*Boreal Ecosystem Atmosphere Study*
BPI	*Bits per Inch*
BPS	*Bits per Second*
BRDF	*Bi-directional Reflection Distribution Function*
BS	*Base Station*
BWC	*Bandwidth Compression*
BWE	*Bandwidth Expansion*
C	
C/A	*Coarse-Acquisition (code)*
CAD	*Computer-Aided Drafting; Computer-Assisted Design; Computer-Aided Design*
CADD	*Computer-Aided Design and Drafting*
CADMAP	*Computer-Aided Drafting, Mapping, and Photogrammetry*
CADRG	*Compressed ADRG*
CALIPSO	*Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations*
CAM	*Computer-Aided Mapping*
CAMA	*Computer-Aided Mass Appraisal System (Montana)*
CAMEO	*Computer-Aided Management of Emergency Operations System (USA)*
CASI	*Compact Airborne Spectrographic Imager (Canada)*
CAT	*Computer-Assisted Thermography*
CBERS	*China–Brazil Earth Resources Satellite Program*
C-CAP	*Coastal Change Analysis Program*
CCD	*Charge-Coupled Device*
CCD/TDI	*Charge-Coupled Device/Time Delay Integration*
CCNS	*Computer Controlled Navigation System (IGI, Germany)*
CCPR	*Consultative Committee on Photometry and Radiometry*
CCRS	*Canada Centre for Remote Sensing (Canada)*
CD	*Change Detection, Compact Disk*
CDM	*Canonical Data Model*
CDR	*CorelDraw format*
CD-R	*Compact Disk-Recordable*
CD-ROM	*Compact Disk-Read Only Memory*
CDS	*Component Database System*
CDTED	*Compressed Digital Terrain Elevation Data*
CD-W	*Compact Disk re-writable*
CENT/TC 287	*Committee of Normalization, Technical Committee 287*
CEO	*European Centre for Earth Observation*
CEOS	*Committee on Earth Observation Satellites*
CERES	*Clouds and Earth's Radiant Energy System (NASA EOS)*
CESBIO	*Centre d' Études Spatiales de la Biosphère (France)*
CGA	*Colour Graphics Adaptor*
CGI	*Common Gateway Interface*
CGM	*Computer Graphics Metafile*
CIE	*Commission Intemationale de l'Éclairage*
CIGNET	*Cooperative International GPS Network*
CILSS	*Comité inter-Etats de Lutte contre la Sécheresse au Sahel*

Acronym	Definition
CIR	*Color InfraRed*
CLC	*CORINE Land Cover*
CMIS	*Conical Scanning Microwave Imager/Sounder*
CML	*Chemical Markup Language*
CMYK	*Cyan–Magenta–Yellow–Black*
CNES	*Centre National d'Études Spatiales (France)*
CNR	*Consiglio Nazionale delle Ricerche (Italy)*
CNRS	*Centre National de la Recherche Scientifique (France)*
COBOL	*COmmon Business-Oriented Language*
CONAE	*Comision National de Actividades Espaciales*
CORINE	*Coordination of Information on Environment*
COSMO/SkyMed	*COnstellation of Small satellites for the Mediterranean basin Observation*
CPU	*Central Processing Unit*
CRSS	*Canadian Remote Sensing Society*
CSA	*Canadian Space Agency*
CSCW	*Computer Supported Cooperative Work*
CSIRO	*Commonwealth Scientific and Industrial Research Organisation (Australia)*
CVA	*Change Vector Analysis*
CZCS	*Coastal Zone Color Scanner*
D	
D/A	*Digital to Analog (converter)*
DAAC	*Distributed Active Archive Center*
DAIS-1	*Digital Airborne Imagery System*
DAM	*Detection and Mapping*
DAT	*Digital Audio Tape*
DATIS	*Digital Airborne Topographic Imaging System*
DB	*DataBase*
DBMS	*DataBase Management System*
DCF	*Digital Cartographic File*
DDR	*Data Descriptive Record*
DDS	*Digital Data Storage*
DDSM	*Dense Digital Surface Model*
DEI	*Department of Electronics and Information*
DEM	*Digital Elevation Model; Digital Elevation Matrix*
DEM-G	*Digital Elevation Model – Graphic (coordinates)*
DEM-P	*Digital Elevation Model – Planar (coordinates)*
DGDF	*Digital Geospatial Data Files*
DGPS	*Differential Global Positioning System; Differential Global Positioning Satellite*
difSAR	*Differential SAR*
DIGCAT	*Digital Catalogue*
DIGEST	*Digital Geographic Information Exchange Standard*
DLL	*Windows Dynamic Link Library*
DLR	*Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Agency)*
DLT	*Digital Linear Tape*
DMA	*Defense Mapping Agency (USA); Designated Market Areas*
DMC	*Disaster Monitoring Constellation*

Acronym	Definition
DN	*Digital Number*
DoD	*Department of Defense (USA)*
DOP	*Dilution of Precision*
DORIS	*Decision-Oriented Resource Information System; Doppler Orbitography and Radiopositioning Integrated by Satellite*
DPI	*Dots per Inch*
DRF	*Digital Raster File*
DSE	*Data Switching Equipment*
DSM	*Digital Surface Model*
DSS	*Decision Support System*
DTD	*Document Type Definition*
DTE	*Data Terminal Equipment*
DTED	*Digital Terrain Elevation Data (U.S. government)*
DTM	*Digital Terrain Model (or Modelling)*
DWF	*Drawing Web Format*
DXF	*Drawing Interchange File; Digital Exchange Format*
E	
EAE	*European Agency for the Environment*
EARSeL	*European Association of Remote Sensing Laboratories*
EarthCARE	*Earth Clouds Aerosols and Radiation Explorer*
EC	*European Commission*
ECSS	*European Cooperation for Space Standardisation*
ED50	*European Datum 1950*
EDC	*EROS Data Center (USGS)*
EE	*Electrical Engineering*
EEA	*European Environment Agency*
EGA	*Enhanced Graphics Adaptor*
EGNOS	*European Geostationary Navigation Overlay Service*
EIA	*Environmental Impact Assessment*
EIONET	*European Environment Information and Observation Network*
ENVISAT	*Environmental Satellite (ESA)*
EO	*Earth Observation*
EO-1	*Earth Orbiter-1 (NASA)*
EOF	*End of File*
EORC	*Earth Observation Research Center (Japan)*
EOS	*Earth Observing System, Earth Observation Satellites*
EOS AM	*NASA EOS Morning Crossing (Descending) Mission*
EOS PM	*NASA EOS Afternoon Crossing (Ascending) Mission*
EOSAT	*Earth Observation Satellite (Company)*
EOSDIS	*EOS Data and Information System (NASA)*
EPA	*Environmental Protection Agency (USA)*
EPS	*EUMETSAT Polar System*
EPS	*Encapsulated PostScript*
ERA	*European Research Area*
ERDAS	*Earth Resources Data Analysis System*
EROS	*Earth Resources Observation Satellite; Earth Resources Observation System*
ERS	*European Remote-Sensing Satellite (ESA)*
ERSAR	*Earth Resources Synthetic Aperture Radar*

Acronym	Definition
ErtPS	*Extended Real-Time Polling Service*
ERTS	*Earth Resources Technology Satellite*
ES	*Expert System*
ESA	*European Space Agency*
ESE	*Earth Science Enterprise*
ESRIN	*European Space Research Institute (ESA establishment)*
ESSA	*Environmental Science Service Administration*
ESSP	*Earth System Science Pathfinder*
ESTB	*EGNOS System Test Bed*
ESTEC	*European Space Research and Space Technology Centre (ESA)*
ETC	*European Topic Centres*
ETM+	*Enhanced Thematic Mapper Plus (NASA Landsat)*
ETRS89	*European Terrestrial Reference System 1989*
EU	*European Union*
EUMETSAT	*European Organisation for the Exploitation of Meteorological Satellites*
EUREF	*European Reference Frame*
F	
FAO	*Food and Agriculture Organization of the United Nations*
FCC	*False Colour Composite; Federal Communications Commission (USA)*
FGDC	*Federal Geographic Data Committee*
FM	*Facilities Management; Field Manual*
FM-1	*Flight Model 1 (NASA EOS MODIS)*
FMC	*Forward Motion Compensation*
FMP	*Field Pack Mobile Professional*
FORTRAN	*Formula Translation (computer language)*
FOV	*Field of View*
FSA	*Federal Space Agency (Russia)*
FTP	*File Transfer Protocol*
G	
G3OS	*Global Terrestrial Observing System*
GALILEO	*Informal name of GNSS*
GATES	*Geostationary Advanced Technology Environmental System*
GB	*Gigabyte; Geologische Bundesanstalt (Austria)*
GBF	*Geographic Base File*
GBI	*Green Biomass Index*
GCOM	*Global Climate Observing System*
GCP	*Ground Control Point(s)*
GCS	*Ground Control Station*
GCT	*Greenwich Civil Time*
GDA	*Geographic Data Analysis System*
GDC	*Geographic Data Council*
GDF	*Geographic Data File*
GEMI	*Global Environment Monitoring MSAVI*
GEO	*Geostationary Orbit*
GEOS	*Geodynamics Experimental Ocean Satellite*
GEOSAT	*Geodesy Satellite*

Acronym	Definition
GeoVIS	*Geographic Vector Interpretation System (Africover)*
GERIS	*Geophysical Environmental Research Imaging Spectrometer*
GGRF	*Galileo Geodetic Reference Frame*
GI	*GeoInformation; GeoSpatial Information*
GIF	*Graphic Interchange Format*
GIS	*Geographic(al) Information System(s); Geohydrologic Information System; Global Indexing System*
GISCO	*Geographic Information System of the European Commission*
GISIG	*Geographical Information Systems International Group*
GLAI	*Green Leaf Area Index*
GLCC	*Global Land Cover Characteristic*
GLCN	*Global Land Cover Network*
GLI	*Global Imager (Japan, ADEOS)*
GLOBE	*Global Change Observing Mission*
GLONASS	*Global'naya Navigatsionnaya Sputnikovaya Sistema, Global Orbiting Navigation Satellite System (Russia)*
GLOSS	*Global Observation Surveillance System*
GLRS	*Geoscience Laser Ranging System*
GMES	*Global Monitoring for Environment and Security*
GML	*Geography Markup Language*
GMS	*Global Land One-km Base Elevation*
GMT	*Greenwich Mean Time*
GNS	*GEOnet Names Server*
GNSS	*Global Navigation Satellite System*
GOCE	*Gravity field and steady-state Ocean Circulation Explorer*
GOES	*Geostationary Operational Environmental Satellite (NASA)*
GOFC	*Global Observations of Forest Cover (IGOS)*
GOLD	*Global Observation of Land Cover Dynamics*
GOME	*Global Ozone Monitoring Experiment*
GOMOS	*Global Ozone Monitoring by Occultation of Stars*
GOMS	*Geostationary Operational Meteorological Satellites*
GPL	*General Public Licence*
GPR	*Ground-Penetrating Radar*
GPRS	*General Packet Radio Service*
GPS	*Global Positioning System(s)*
GPST	*Global Positioning System Time*
GRASS	*Geographic Resource Analysis Support System*
GRD	*Ground Resolved Distance*
GRS 80	*Geodetic Reference System of 1980*
GSD	*Ground Sample Distance*
GSDI	*Global Spatial Data Infrastructure*
GSDM	*Global Spatial Data Model*
GSDS	*Global spatial data system*
GSFC	*Goddard Space Flight Center (NASA)*
GSLV	*Geosynchronous Satellite Launch Vehicle*
GSM	*Global System for Mobile Communication – Graphic Standard Metafile*
GSRC	*Geological Survey Research Committee*
GSS	*Geographic Support System*
GST	*Galileo System Time*

Acronym	Definition
GSTB	*Galileo System Test Bed*
GTOS	*Global Terrestrial Observing System (FAO)*
GUI	*Graphical User Interface*
GVI	*Green Vegetation Index*
H	
HIPERLAN	*High-PErformance Radio LAN*
HIRDLS	*High-Resolution Dynamics Limb Sounder*
HIRIS	*Hyperspectral InfraRed Imaging Spectrometer*
HPGN	*High-Precision Geodetic Network*
HRC	*High-Resolution Camera*
HRG	*High-Resolution Geometry (SPOT)*
HRPC	*High-Resolution Panchromatic Camera*
HRSC-A	*High-Resolution Stereo Camera-Airborne*
HRV	*Haute Résolution dans le Visible (SPOT)*
HRVIR	*High Resolution Visible Infrared (scanner)*
HSI	*Hue, Saturation, Intensity – Hyperspectral Imaging Instrument*
HTML	*Hypertext (or HyperText) Markup Language*
HTTP	*HyperText Transfer Protocol*
HTV	*H-II Transfer Vehicle*
HW	*Hardware*
HYDICE	*HYperspectral Digital Imagery Collection Experiment*
HYMAP	*Airborne Hyperspectral Imaging*
I	
I&CLC2000	*Image and Corine Land Cover 2000*
I/O	*Input/Output*
IASI	*Infrared Atmospheric Sounding Interferometer*
ICAO	*International Civil Aviation Organisation*
ICC	*International Colour Consortium*
IDC	*Internet Data Center*
IDE	*Imbedded Drive Electronics*
IDL	*Interactive Data Language; Interface Definition Language*
IDS	*Interdisciplinary Science*
IEC	*International Electrotechnical Commission*
IERS	*International Earth Rotation Service*
IFAD	*International Fund for Agricultural Development*
IFOV	*Instantaneous Field of View*
IFSARE	*Inter-ferometric Synthetic Aperture Radar for Elevation*
IGADD	*Intergovernmental Authority on Drought and Development*
IGBP	*International Geosphere–Biosphere Programme*
IGBP-DIS	*International Geosphere–Biosphere Programme – Data and Information System*
IGOS	*Integrated Global Observing Strategy*
IGS	*International GPS Service*
IKAR-N	*scanning microwave radiometer system*
IKAR-P	*nadir microwave radiometer*
ILAS	*Improved Limb Atmospheric Spectrometer (Japan)*
IMG	*Interferometric Monitor for Greenhouse Gases*
IMO	*International Maritime Organisation*

Acronym	Definition
IMS	*Information Management System*
IMU	*Inertial Measurement Unit*
INPE	*Instituto Nacional de Pesquisas Espaciais (Brasil)*
INRA	*Institut National de la Recherche Agronomique (France)*
INS	*Inertial Navigation System*
InSAR	*Interferometric Synthetic Aperture Radar*
INSAT	*Indian National Satellite*
INSPIRE	*Infrastructure of Spatial Information in the European Community*
IP	*Internet Protocol – Image Point – Integrated Project (CE)*
IPA	*Imagery Product Archive (USDMA)*
IPS	*Image Processing System; Inertial Positioning System*
IR	*Imaging Radiometer; Infrared; Information Retrieval*
IRF	*Internal Raster File; Internal Raster Format*
IRMSS	*Infrared Multispectral Scanner*
IRS	*Internal Revenue Service (USA); Information Retrieval System; Indian Remote Sensing (Satellite)*
ISA	*Israelian Space Agency*
ISAMS	*Improved Stratospheric and Mesospheric Sounder*
ISAS	*Institute of Space and Astronautical Science*
ISO	*International Standards Organization*
ISPRS	*International Society for Photogrammetry and Remote Sensing*
ISRO	*Indian Space Research Organisation*
ISS	*International Space Station*
ITOS	*Improved Tiros Operational Satellite*
ITRF	*International Terrestrial Reference Frame*
ITRS	*International Terrestrial Rotation Service*
ITU	*International Telecommunication System*
J	
JAVA	*Joint Academic Virtual Application*
JAXA	*Japan Aerospace Exploration Agency (ex NASDA)*
JEM	*Japan Experiment Module*
JERS	*Japanese Earth Resources Satellite*
JFIF	*JPEG File Interchange Format*
JGW	*JPEG with World File*
JPEG	*Joint Photographic Experts Group format*
JPL	*Jet Propulsion Laboratory (USA)*
JRC	*Joint Research Centre (Italy)*
JSP	*Java Server Pages*
JVM	*Java Virtual Machine*
K	
KARI	*Korea Aerospace Research Institute*
KB	*Kilobyte*
KBPS	*KiloBytes Per Second*
KOMPSAT	*Korean Multipurpose Satellite*
L	
LADAR	*Laser Detection and Ranging*
LFC	*Large Format Camera*

Acronym	Definition
LAGEO	*Laser Geodynamic Satellite*
LAI	*Leaf Area Index*
LAMP	*Low Altitude Mapping Photogrammetry*
LAMPS	*Laser-scan Automated Map Production System*
LAN	*Local Area Network*
LANDSAT	*Land Satellite*
LARA	*Aerial Laboratory for Environmental Research*
LASER	*Light Amplification by the Stimulated Emission of Radiation*
LBS	*Location-Based Services*
LCA	*Land Cover analysis*
LCCP	*National Land Cover Characterization Project*
LCCS	*Land Cover Classification System*
LCD	*Liquid-Crystal Displays*
LCDB	*Land Cover Database*
LCS	*Location Services*
LDHF	*Landsat-7 Data Handling Facility*
LEO	*Low Earth Orbit*
LGPL	*General Public Licence Library*
LIDAR	*Light Detection and Ranging*
LIDQA	*Landsat Image Data Quality Assessment (NASA)*
LIS	*Land Information System*
LISS	*Linear Imaging Self-scanning Sensor*
LORAN	*Long Range Navigation (system)*
LPI	*Lines per inch*
LPV	*Land Product Validation*
LRR	*Laser Reflector*
LRS	*Landsat Recorder System*
LSPIM	*Land-Surface Processes and Interactions Mission*
LSS	*Laser Scanning System*
LUT	*Look Up Table*
LWIR	*Long-Wave Infrared (spectral region)*
LZW	*Lempel Ziv Welch*
M	
MacOS	*Machintosh Operative System*
MADE	*Multipurpose Africover Database for Environmental resources*
MAGSAT	*Magnetic Field Satellite (USA)*
MAIS	*Modular Airborne Imaging Spectrometer*
MAP	*Map Accuracy Programme (Africover)*
MAPSAT	*Mapping Satellite (USA)*
MAS	*MODIS Airborne Simulator (NASA EOS)*
MATLAB	*Matrix Laboratory*
MB	*Megabyte*
MBLA	*Multi-Beam Laser Altimeter*
MCDM	*Multi-Criteria decision making*
MCGA	*Monochrome/Colour Graphics Adapter*
MCST	*MODIS Characterization Support Team (NASA EOS)*
MD	*Minimum Distance*
MEBS	*MODIS Emergency Backup System (NASA EOS)*
MEIS	*Multispectral Electrooptical Imaging Sensor*

Acronym	Definition
MEO	*Medium Earth Orbit*
MERES	*Mineral and Energy Resources Exploration Satellite (Japan)*
MERIS	*Medium Resolution Imaging Spectrometer (ESA Envisat)*
MESSR	*Multispectral Electronic Self-Scanning Radiometer*
METOP	*METeorology OPerational*
METSAT	*Meteorological Satellite*
MGRS	*Military Grid Reference System*
MHS	*Microwave Humidity Sounder (NASA EOS)*
MHz	*MegaHertz*
MIAS	*Multispectral Image Analysis System*
MIDAS	*Multispectral Interactive Data Analysis System*
MIME	*Multipurpose Internet Mail Extension*
MIPAS	*Michelson Interferometer for Passive Atmospheric Sounding*
MIR	*Russian orbital station*
MIRAS	*Microwave Imaging Radiometer with Aperture Synthesis (ESA)*
MISR	*Multiangle Imaging Spectro-Radiometer (NASA EOS)*
MIT	*Massachusetts Institute of Technology*
MIVIS	*Multispectral Infrared and Visible Imaging Spectrometer*
MLA	*Multi-linear array*
MLL	*Maximum Likelihood*
MLP	*Multilayer Perceptron*
MLS	*Microwave Limb Sounder*
MNS	*Multi-Navigator System*
MODEM	*Modulation and De-Modulation*
MODIS	*Moderate-Resolution Imaging Spectroradiometer (NASA EOS)*
MODLAND	*MODIS Land Discipline Group (NASA EOS)*
MODS	*Moderate Resolution Imaging Spectrometer*
MODEM	*MOdulation and DEModulation*
MOMS	*Modular Optoelectric Multispectral Scanner*
MOPITT	*Measurements of Pollution in the Troposphere (Canada; NASA EOS)*
MP-MP	*Multipoint–Multipoint Systems*
MPRD	*Modified Proportional Radial Displacement*
MrSID	*Multiresolution Seamless Image Database*
MS	*MultiSpectral; Microwave Sounder*
MSAS	*MTSAT Satellite-Based Augmentation System*
MSAVI	*Modified Soil Adjusted Vegetation Index*
MSC	*Multispectral Camera*
MSG	*Meteosat Second Generation (ESA)*
MSL	*Mean Sea Level*
MSMR	*Multi-frequency Scanning Microwave Radiometer*
MSP	*Microsoft Paint format*
MSR	*Microwave Scanning Radiometer*
MSS	*Multispectral Scanner (Landsat)*
MSU-SK	*Multispectral Optical Scanner*
MTF	*Modulation Transfer Function*
MTPE	*Mission to Planet Earth (NASA)*
MTSAT	*Multi-functional Transport Satellite (Japan)*
MTU	*Magnetic Tape Unit*
MWF	*Map Window File*

Acronym	Definition
N	
NAD	*North American Datum*
NAL	*National Aerospace Laboratory of Japan*
NAPA	*National Academy of Public Administration*
NARSA	*National Advanced Remote Sensing Applications Program*
NASA	*National Aeronautics and Space Administration (USA)*
NASDA	*National Space Development Agency (Japan)*
NATO	*North Atlantic Treaty Organization*
NAVSTAR	*NAVigation Satellite Timing and Ranging Global Positioning System*
NAVTECH	*Navigation Technologies Corp.*
NBSM	*National Bureau of Surveying and Mapping (China)*
NCC	*National Cartographic Centre (Iran)*
NDCS	*National Digital Cartographic Standards*
NDVI	*Normalized Difference Vegetation Index*
NERC	*Natural Environment Research Council (UK)*
NESDIS	*National Space Development Agency of Japan*
NET	*National Environmental Satellite, Data, and Information Service*
NGDC	*National Geophysical Data Center (USA); National Geospatial Data Clearinghouse*
NGDF	*National Geospatial Data Files*
NGMA	*National Geoscience Mapping Agency (Australia)*
NGMBD	*National Geologic Map Database (USGS)*
NGWIC	*National Ground Water Information Center*
NICMOS	*Near-Infrared Camera and Multi-Object Spectrometer*
NIMA	*National Imagery and Mapping Agency (USA)*
NIR	*Near Infrared (spectral region)*
NLAPS	*National Landsat Archive Production System (USGS)*
NLCD	*National Land Cover Dataset (USGS),*
NLOS	*Not Line Of Sight*
NMAS	*National Map Accuracy Standard(s)*
NMP	*New Millenium Program (NASA)*
NNSS	*Navy Navigational Satellite System*
NOAA	*National Oceanic and Atmospheric Administration (USA)*
NPOESS	*National Polar-orbiting Operational Environmental Satellite System (USA)*
NRSA	*National Remote Sensing Agency (India)*
NRSC	*National Remote Sensing Centre (UK)*
NRSCC	*National Remote Sensing Centre of China*
NSAU	*National Space Agency of Ukraine*
NSC	*Norwegian Space Centre*
NSCAT	*NASA Scatterometer*
NSRS	*National Spatial Reference System (NOAA)*
NSP	*Korean National Space Development Plan*
NTF	*National Transfer Format (UK)*
NTIS	*National Technical Information Service (USA)*
NTS	*National Topographic Series (Canada); National Topographic System (Canada); Navigation Technology Satellite*
NVCS	*National Vegetation Classification Standard*

Acronym	Definition
O	
OCM	*Ocean Colour Monitor*
OCR	*Optical Character Recognition; Optical Character Reader*
OCTS	*Ocean Colour and Temperature Scanner (Japan, ADEOS)*
ODBC	*Operational Database Connectivity*
OEC	*Optical-Electronic Camera*
OGC	*Open Geospatial Consortium*
OLE	*Object Linking Embedding*
OMI	*Ozone Monitoring Instrument*
ORBIMAGE	*Orbital Imaging Corp*
ORD	*Object-Relational DataBase*
ORDBMS	*Object-Relational DataBase Management System*
ORFEO	*Optical and Radar Federated Earth Observation*
OS	*Operative System*
OSCAR	*Online Satellite Catalogue and Request System (NOAA)*
OSMI	*Ocean Scanning Multispectral Imager*
OSTM	*Ocean Surface Topography Mission*
P	
PAGES	*Program for the Adjustment of GPS Ephemeris*
PALSAR	*Phased-Array L-band SAR (Japan)*
PAN	*Panchromatic*
PBM	*Portable Bitmap (format)*
PC	*Personal Computer; Principal Components*
PCA	*Principal Components Analysis*
P-Code	*Precise Code; Protected Code*
PCX	*PC Paintbrush Export Format*
PDA	*Personal Digital Assistant*
PE&RS	*Photogrammetric Engineering & Remote Sensing (journal)*
PFM	*Proto-Flight Model (NASA)*
PGM	*Portable Graymap (format)*
PHP	*Personal Home Page*
PIC	*Lotus 1-2-3 Picture Format*
Pixel	*Picture element*
P-MP	*Point–MultiPoint systems*
PNG	*Portable Network Graphics (format)*
PNN	*Probabilistic Neural Network*
POLDER	*Polarization and Directionality of the Earth's Reflectances (CNES, ADEOS)*
PP	*Principal Point*
PPM	*Portable Pixelmap (format)*
PPR	*Portland Pattern Repository*
PPS	*Precise Position Service*
PR	*Precipitation Radar (Japan)*
PRARE	*Precise Range and Range Rate Equipment*
PRF	*Point Response Function*
PRIRODA	*Nature Instrument Model on Russian MIR Station*
PRISM	*Panchromatic Remote Sensing Instrument for Stereo Mapping (Japan); Processes Research by an Imaging Space Mission*

Acronym	Definition
PRS	*Public Regulated Service*
PSLV	*Polar Satellite Launch Vehicle*
PVI	*Perpendicular Vegetation Index*
Q	
QA	*Quality Assurance*
QUASAR	*QUality Assurance and StAbility Reference*
R	
RA	*Radar Altimeter*
RADAR	*Radio Detecting and Ranging*
RAM	*Random Access Memory*
RAR	*Real Aperture Radar*
RASA	*Russia Aviation-Space Agency*
RBV	*Return Beam Vidicon*
RCSSMRS	*Regional Center for Services in Surveying, Mapping and Remote Sensing*
RDBMS	*Relational Database Management System*
RFM	*Rational Functional Model*
RGB	*Red–Green–Blue*
RIN	*Realty Information Network*
RINEX	*Receiver Independent Exchange (GPS data format)*
RIP	*Raster Image Processing; Raster Image Processor*
RIS	*Retroreflector in Space*
RMSE	*Root-Mean-Square Error*
RNSS	*Radio Navigation Satellite Service*
ROA	*Remotely Operated Aircraft*
ROM	*Read-Only Memory*
ROSA	*Radio Occultation for Sounding the Atmosphere*
ROSHYDROMET	*Russian Federal Service for Hydrometeorology and Environmental Monitoring*
ROSIS	*Reflective Optics System Imaging Spectrometer*
RPC	*Rational Polynomial Coefficients*
RPF	*Raster Product Format*
RPV	*Remotely Piloted Vehicle*
RSA	*Russian Space Agency Rosaviakosmos*
RSG	*Remote Sensing Group (UAZ)*
RSI	*RADARSAT International*
RSS	*Remote Sensing Society (UK)*
RTCM	*Radio Technical Commission for Maritime (format)*
RTF	*Rich Text Format*
RT-GPS	*Real-Time Differential GPS*
rtPS	*Real-Time Polling Service*
RTK	*Real-Time Kinematics; Right-to-Know Network*
RTP	*Real-Time Positioning*
S	
S/N	*Signal-to-Noise Ratio*
SA	*Selected Availability; Selective Availability*
SADCC	*Southern African Development Community*

Acronym	Definition
SAN	*Storage Area Network*
SAR	*Synthetic Aperture Radar; Search and Rescue Service*
SAVI	*Soil-Adjusted Vegetation Index*
ScaLARS	*Scanning Laser Altitude and Reflectance Sensor (Germany)*
SCARAB	*Scanner for the Radiation Budget*
SCIAMACHY	*Scanning Imaging Absorption Spectrometer for Atmospheric Cartography*
SDBMS	*Spatial DataBase Management System*
SDES	*Spatial Data Exchange Standard*
SDI	*Spatial Data Infrastructure*
SE	*Search Engine*
SEASAT	*Sea Satellite (USA)*
SeaWiFS	*Sea-Viewing Wide Field-of-View Sensor*
SEVIRI	*Spinning Enhanced Visible and Infrared Imager (ESA)*
SFSI	*Shortwave-Infrared Full Spectrum Imager (Canada)*
SI	*Spatial Information*
SIASGE	*Sistema Italo-Argentino de Satélites protects Gestión de Emergencias*
SIR	*Shuttle Imaging Radar; Space Imaging Radar (NASA)*
SIRS	*Satellite Infrared Spectrometer*
SIS	*Spatial Information System*
SLAR	*Side-Looking Airborne Radar*
SMIL	*Synchronized Multimedia Integration Language*
SMOS	*Soil Moisture and Ocean Salinity*
SMTP	*Simple Mail Transfer Protocol*
SNM	*Square Nautical Miles*
SNN	*Structured Neural Network*
SNSB	*Swedish National Space Board*
SoL	*Safety of Life Service*
SOLSPEC	*Solar Spectrum (Space Shuttle instrument from Europe)*
SOM	*Space Oblique Mercator (projection); Spatial Object Manager*
SONAR	*Sound Navigation and Ranging*
SPIE	*Society of Photo-optical Instrumentation Engineers*
SPIN-2	*Space Information-2 Meter Data (Russian Space Agency)*
SPOT	*Systeme Probatoire d'Observation de la Terre; Systeme Pour l'Observation de la Terre*
SPS	*Standard Position Service*
SQL	*Standard Query Language; Structured Query Language*
SRMCS	*Soil Resources Management and Conservation Service*
SROM	*Spectro-radiometer for Ocean Monitoring*
SRTM	*Shuttle Radar Topography Mission*
SS	*Subscriber Station*
STDN	*Spacecraft Tracking and Data Network*
STS	*Space Transportation System*
SVG	*Scalable Vector Graphics*
S-VGA	*Super Video Graphics Array*
SW	*Software*
SWIR	*Short-Wave InfraRed (TM5 and TM7 spectral region)*
T	
TCP/IP	*Transmission Control Protocol/Internet Protocol*
TCT	*Tasseled Cap Transformation, Kauth–Thomas 1976*

Acronym	Definition
TDRSS	*Tracking and Data Relay Satellite System (NASA)*
TES	*Tropospheric Emission Spectrometer*
3S	*Small Satellite System, Suite de Systeme SPOT, SPOT Successor System*
TIFF	*Tag Image File Format*
TIMS	*Therms Infrared Multispectral Scanner*
TIN	*Triangulated Irregular Network*
TIR	*Thermal InfraRed*
TIROS	*Television Infrared Observing Satellite*
TIRS	*Thermal Infrared Scanner*
TIS	*Topographic Information Systems*
TLS	*Terrestrial Laser Scanning*
TM	*Thematic Mapper (Landsat) – Telemetry*
TMI	*TRMM Microwave Instrument (Japan)*
TNC	*The Nature Conservancy*
TOMS	*Total Ozone Mapping Spectrometer*
TOPEX	*Ocean Topography Experiment*
TP	*Tie Point*
TPS	*Thin-Plate Splines*
TRMM	*Tropical Rainfall Measuring Mission (Japan)*
TSAVI	*Transformed Soil Adjusted Vegetation Index*
TT&C	*Tracking, Telemetry and Command*
U	
U-2	*A high altitude remote sensing aircraft*
UARS	*Upper Atmosphere Research Satellite*
UAV	*Unmanned Aerial Vehicles*
UGS	*Unsolicited Grant Service*
UI	*Urban Index*
UID	*Univariate Image Differencing*
UMTS	*Universal Mobile Telecommunications System*
UNEP	*United Nations Environment Program*
UNOOSA	*United Nations Office of Outer Space Affairs*
UN-WFP	*United Nations-World Food Programme*
URI	*Uniform Resource Identifier*
URL	*Uniform Resource Locator*
USDA	*United States Department of Agriculture*
USGS	*United States Geological Survey (USA)*
USNVC	*United States National Vegetation Classification*
UTC	*Universal Time Coordinated*
UTM	*Universal Transverse Mercator*
UTP	*Unshielded Twisted-Pair*
UV	*Ultraviolet*
UVS	*Unmanned Vehicle System*
V	
VCL	*Vegetation Canopy Lidar*
VGA	*Video Graphics Array; Video Graphics Adaptor*
VGIS	*Virtual Geographical Information System*
VGT	*SPOT-4 VEGETATION instrument (France)*
VHF	*Very High Frequency*

Acronym	Definition
VHRPC	*Very High Resolution Panchromatic Camera*
VHRR	*Very High Resolution Radiometer*
VI	*Vegetation Index*
VID	*Vegetation Index Differencing*
VIIRS	*Visible and Infrared Imaging Radiometer Suite*
VIRS	*Visible Infrared Scanner (Japan)*
VIS	*Visible (spectral region)*
VLDS	*Very Large Data Store*
VML	*Vector Markup Language*
VNIR	*Visible and Near-Infrared (spectral region)*
VoIP	*Voice over Internet Protocol*
VPN	*Virtual Private Networks*
VTIR	*Visible and Thermal Infrared Radiometer*
W	
W3C	*World Wide Web Consortium*
WAAS	*Wide Area Augmentation System*
WAN	*Wide Area Network*
WAP	*Wireless Application Protocol*
WCRP	*World Climate Research Program*
WDVI	*Weighted Difference Vegetation Index*
WFI	*Wide Field Imager*
WGCV	*Working Group on Calibration and Validation (CEOS)*
WGS 84	*World Geodetic System of 1984*
Wi-Fi	*Wireless Fidelity*
WiFS	*Wide Field-of-view Sensor*
WiMAX	*Worldwide Interoperability for Microwave Access*
WirelessMAN	*Wireless Metropolitan Area Network*
WIS	*West Indian Space*
WLAN	*Wireless Local Area Network*
WMO	*World Meteorological Organization*
WWW	*World Wide Web*
X	
xDLS	*eXtensible Distance Learning System*
XML	*eXtensible Mark-up Language*
XSL	*eXtensible Stylesheet Language*
XSLT	*eXtensible Stylesheet Language Transformation*

Chapter 1
Geomatics

. . . for thousands of years every man had counted only his own territory but the measure of the terrestrial circumference will mark for the men the exit from the village; and every man, transcending the connection to his land, will become an inhabitant of the Earth. . .

Eratosthenes (275–193 BCE)
Greek mathematician–astronomer–geographer–poet

Man has always wished to explore, to know and to represent the places where he lives. Tracing itineraries to reach various places and identifying favourable sites for agricultural use were the beginnings of geography, which then became a science in the XVI century thanks to the talent and foresight of Leonardo da Vinci.

Geography, through network systems, data-flow analysis, knowledge of economic factors and the socio-political organization of the territory, has now become an independent discipline supplying a synthetic vision of our planet and of the complex relations between physical and man-induced phenomena.

From classical geography, scientific activities in Earth Observation have undergone a rapid expansion, and more and more economic sectors tend to employ territorial data acquired by ground survey, global satellite positioning systems, traditional and digital photogrammetry, multi- and hyperspectral remote sensing from airplane and satellite, with images both passive optical and active microwave (radar) at different geometric, spectral, radiometric and temporal resolutions, although there is still only limited awareness of how to use all the available potential correctly. The resulting data and information are represented in digital and numerical layers managed in Geographical Information Systems and Decision Support Systems, often based on the development of Expert Systems.

Such a large amount of data must necessarily be organized, processed, handled and used without delay for a correct representation of the territorial situation.

These elements must be processed in an interdisciplinary and interoperable manner, and the discipline of geomatics (*geos*: Earth, *matics*: informatics) can satisfy such requirements.

The term *geomatics* was created at Laval University in Canada in the early 1980s, based on the concept that the increasing potential of electronic computing was

M.A. Gomarasca, *Basics of Geomatics*, DOI 10.1007/978-1-4020-9014-1_1,

revolutionizing surveys and representation sciences and that the use of computerized design (video-diagram) was compatible with the treatment of huge amounts of data. That period's revolutionary intuition was based on the geographical location of each object on our planet.

> Geomatics is defined as a systemic, multidisciplinary, integrated approach to selecting the instruments and the appropriate techniques for collecting, storing, integrating, modelling, analysing, retrieving at will, transforming, displaying and distributing spatially georeferenced data from different sources with well-defined accuracy characteristics, continuity and in a digital format.

Erected on the scientific framework of geodesy, it uses terrestrial, marine, airborne and satellite-based sensors to acquire spatial and other data.

Some initiatives are presently developing worldwide using geomatics disciplines and techniques for the regulation of *GeoSpatial Information*, or more simply *GeoInformation* (GI) and for the adequate use of *Earth Observation* (EO) data for studying and managing environmental hazards and risks.

Several countries, following common fundamental guidelines and procedures, are developing a *Spatial Data Infrastructure* (SDI). At planet level, the dream is to realize a *Global Spatial Data Infrastructure (GSDI)* capable of managing heterogeneous sets of data and to overcome the chronic absence of interoperability among *databases* (DB). This goal can be reached by implementing several SDIs at local, national, continental level and harmonizing them in a global context.

One practical example is represented by the *Infrastructure for Spatial Information in the European Community* (INSPIRE), a European Union Directive that entered into force on May 15, 2007.

Other initiatives promote the collective effort for a better Earth environment, by increasing our understanding of the Earth's dynamic processes and enhancing forecasts of our environmental conditions. The *Group on Earth Observations* (GEO) was formed to undertake this global effort, and the *Global Earth Observation System of Systems* (GEOSS) was established on February 16, 2005, with the scope of addressing all nations involved to produce and to manage their information in a way that benefits the environment as well as humanity by taking the pulse of the planet.

As an example, the European initiative GMES (*Global Monitoring for Environment and Security*) is intended to propose solutions for an articulate, centrally coordinated system for risk management at a European level, contributing to GEOSS.

The disciplines and techniques constituting geomatics are

- *Computer science*: to represent and process applicable information through the development of technological instruments (i.e. hardware) and of methods, models and systems (i.e. software).
- *Geodesy*: to determine the shape and size of the Earth; it defines on the one hand the surface of reference in its complete form, the geoid, as well as in its simplified form, the ellipsoid, and on the other hand the external gravitational field as a function of time.

- *Topography*: started with and part of geodesy, this is a combination of procedures for direct land survey. Topography is a combination of methods and instruments to comprehensively measure and represent details of the Earth's surface:

  - *planimetry*: to determine the relative positions of the representation of points on the Earth's surface with respect to the same reference surface;
  - *altimetry*: to determine the height of the points on the Earth's surface with respect to the geoid surface;
  - *tachymetry*: for the planimetric and altimetric survey of the Earth's surface zones;
  - *land surveying*: to measure areas, moving and rectify borders, levelling zones of the Earth physical surface.

- *Cartography*: to supply a possible description of the shape and dimension of the Earth and its natural and artificial details, by means of graphical or numerical representation of more or less wide areas, following fixed rules.
- *Photogrammetry*: to determine the position and shapes of the objects by measuring them on photographic images.
- *Remote Sensing*: to remotely acquire territorial and environmental data and to combine methods and techniques for subsequent processing and interpretation (this definition also fits digital photogrammetry).
- *Global Positioning System (GPS)*: to provide the three-dimensional (3D) position of fixed or moving objects, in space and time, all over the Earth's surface, under any meteorological conditions and in real time.
- *Laser scanning system*: to locate objects and measure their distance by means of the incident radiation in the optical frequencies (0.3–15 μm) of the electromagnetic spectrum.
- *Geographical Information System* (GIS): to make use of a powerful combination of instruments capable of receiving, recording, recalling, transforming, representing and processing georeferenced spatial data.
- *Decision Support System* (DSS): to implement complex Geographical Information Systems, meant to create possible scenarios by modelling the ground truth and to offer a set of solutions to the decision maker.
- *Expert System (ES)*: to consider instruments capable of imitating the experts' cognitive processes and their ability to manage the complexity of reality by means of interdependent processes of abstraction, generalization and approximation.
- *WebGIS*: to distribute geographic data remotely stored on dedicated machines for databases, according to complex network architectures.
- *Ontology*: to specify a conceptuality, i.e. the description of concepts and relationships existing for an element or among various elements of a group, entity or class; conceptualization is an abstract simplified vision of the world to be represented for a given application.

1.1 Computer Science

Computer science culture is now widespread, not only as regards technical instruments but also as regards methods, models and systems that contribute improving activities and research. Informatics, as a discipline, comprises both the computer technologies, i.e. hardware, or the physical components, and software, that is to say the way in which information is structured and elaborated.

The important role of information derives from man's necessity to manage more and more numerous and complex data in every field. The scientific aspect of informatics is emphasized by the systematic, rigorous elaboration of information and its increasingly sophisticated automation.

The human brain constantly elaborates data through both simple and complex cognitive processes, often without any apparent effort. Every action is the result of a more or less complex elaboration by inputting continuously updated data, which are elaborated or simply stored for subsequent processing and then released as output data, or information.

A close relationship exists between the possibility of data representation and data processing; human thinking is a form of calculation, and our mind outlines the truth through symbols and ideas. When a problem exceeds a certain degree of complexity and speed is of the essence, then human elaboration processes become problematic and, in some cases, cannot even be implemented.

The *Association for Computing Machinery* (ACM), an organization for informatics researchers and professionals, has issued the following definition, to underline the planning and developing stages in any informatics activity:

> ...systematic study of the algorithms which describe and transform information: their theory, analysis, plan, efficiency, realization and application...

Algorithms are precise sequences of comprehensible operations, performable by an automatic instrument. Systems for automatic data elaboration are the result of a long process of evolution of elementary instruments for mechanical storage, invented by humans, in order to simplify the performance of calculations, from abacuses to the more recent mechanical calculators.

Current general-purpose devices may store sets of data, operations and instructions, and they may perform such operations in succession, independently by working on data. Programs define the sequence of operations to be performed by expressing them in a language interpretable by the calculator.

1.2 Data and Information

The two terms *data* and *information* are often used synonymously, while their individual meaning is in reality deeply different. The substantial distinction between them is comparable to the difference between an apparently disorganized set of letters and a word assembled with the same letters. Data are the basis of information

and in general represent the measure of the external world. Only an expert system (whether human or not) is able to convert data into information, by reading it according to established rules. The acquisition of information goes via a cognitive process based on data.

In regard to geomatics, the following are a few possible examples:

- a raw satellite image represents the data and any finalized elaboration of it generates information (thematic images, topographical maps);
- Global Positioning System (GPS) data are a measure of time, the derived information is a position in space;
- records of an electronic archive or database are data, the answer to a query using data in a database generates information.

1.3 Geodesy and Cartography

Representing the Earth in a synthetic and exhaustive way and as accurately as possible has been a great challenge for researchers since ancient times. In various ways and in line with the knowledge of the time, historical documentation presents geographic and territorial information using different media and picturesque techniques.

The maps are numerical–graphical products, where the measure and understanding of a territory is synthetically reported. Nowadays, they are generally assembled with restitution procedures of aerial photogrammetry, so providing a representation of the territory at a defined reduction scale, including planimetric details and shapes, and altimetric attributes.

Synthetically speaking, cartography can be defined as the representation of the Earth's surface based on specific rules. The International Association of Cartography (the world leading institution for cartography, defined as a discipline dealing with the conception, production, distribution and study of maps) has supplied the following definition:

> A map is a symbolised image of geographic reality, representing selected features or characteristics, resulting from the creative effort of its author's choices, and it is designed for use when spatial relationships are of primary relevance (Sept. 1995).

Cartography takes advantage of traditional domains such as physics, geometry, design, geography, engineering, with the addition, during the last century, of methodological statistics and electronic numerical calculation for the rigorous elaboration of the data.

Cartography also has a direct relationship with other disciplines for measuring and representing the physical surface of our planet 'from its extensive complexity to its smaller detail': geodesy and topography. The relationships among these disciplines are shown in Fig. 1.1.

Geodesy is the science which defines the shape and dimensions of the Earth through its two branches:

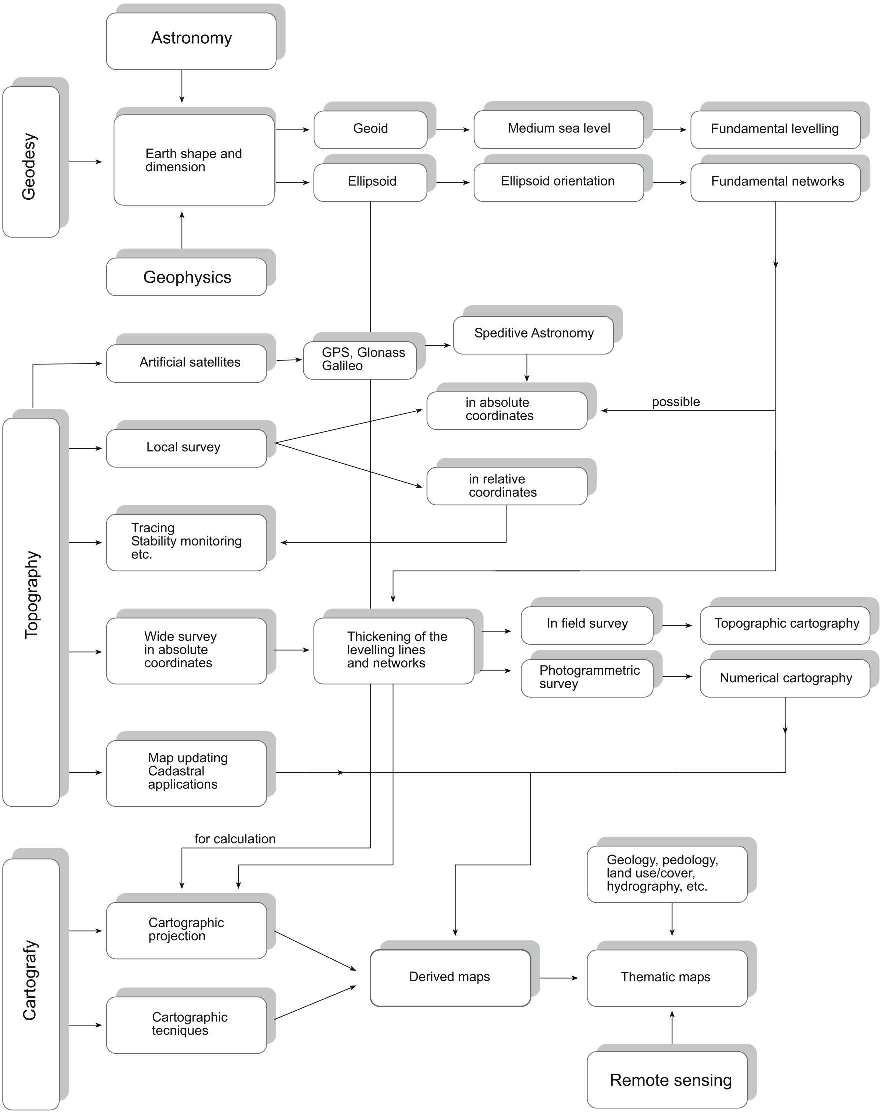


Fig. 1.1 Relationship among the most relevant elements of geomatics

- *gravimetry*, for the determination of the Earth's gravity and its anomalies; the Earth gravity determines the shape of the Earth, and the geoid is the equipotential surface of the gravitational field;
- *positioning astronomy*, for determining the position of the points on the globe through the observation of stars and artificial satellites, referring to the laws of celestial mechanics.

Topography (*topo:* place, *graphia*: writing) deals with the graphical and metric description of sites on the Earth's physical surface features. This discipline was defined as practical geometry until the 18th century, to indicate its operating characteristics of survey and subsequent representation of the territory.

Photogrammetry, introduced in the 19th century, and remote sensing, introduced since the early 1970s, complete the picture.

Attention should be drawn to the use of the terms *representation* and *projection* of the planet Earth, which is a complex system that cannot be replaced with a geometric figure. It would be more correct to say *representation* on a plane of the reference surface rather than *projection* of the Earth. The *representation* of curved surfaces on a plane involves stretching, shrinking and tearing, resulting in interruptions.

As a matter of fact, by applying the simple laws of projective geometry, an area of the Earth's surface is considered projected from a particular point of view using the three basic projection surfaces, plane, cylinder or cone, in the so-called geometric projections.

1.4 Photogrammetry (Analogical, Analytical, Digital)

Photogrammetry is defined as the process of deriving metric information about an object through measurements of the object made on photographs, leaving to photo-interpretation (by human visual analysis) the task to obtain qualitative information (human experience remains a determinant factor).

With the introduction of technologies for image data acquisition from space in a wider region of the electromagnetic spectrum, the meaning of photogrammetry and photo-interpretation has been extended to comprise remote sensing, moving from classical photo-interpretation to the use of digital image processing in addition to human interpretation, and applying computer analysis techniques to imagery besides photography.

A widely accepted definition of photogrammetry and remote sensing is the following:

> Art, science and technology to obtain valid information about physical objects and the environment, through the processes of collection, measure and interpretation of images (photographic or digital) and analog or digital representation of the models of electromagnetic energy derived from survey systems (photographic cameras or scanning systems), without contact with the objects.

Traditional and digital photogrammetries are both based on the same fundamental principles. The tendency to use photogrammetry with high-resolution satellite data invites us to reconsider the traditional approach of the central perspective, while moving its formalisms towards more complex projective geometries that are more connected with the acquisition methods of satellite imagery and new aerophotogrammetric digital cameras. The rigorous reconstruction of the geometric correspondence between image and object at the moment of acquisition remains the primary scope of photogrammetry.

1.5 Remote Sensing

Remote sensing includes techniques to derive information from a site at a known distance from the sensor. In passive remote sensing, the source of information is scattered and/or absorbed solar and emitted thermal radiation, which allows us to study and characterize objects through their spectrally variable response. Every element on the Earth reflects, absorbs and transmits part of an incident radiation to different percentages according to its structural, chemical and chromatic qualities. Electromagnetic radiance is the information reaching a sensor from the objects located on the Earth.

If in passive remote sensing the Sun is the most familiar source of energy, in active remote sensing the sensor itself is at the same time emitter and receiver of electromagnetic energy. The principle used by the latter is the RADAR (*Radio Detection And Ranging*) effect, and the palindrome well portrays the idea of emission and successive recording of the returning signal (*backscattering*) for determining the distance by means of electromagnetic waves with wavelength between 1 mm and 1 m (microwaves).

In common parlance, Remote Sensing, *Télédétection* (French), *Fernerkundung* (German), *Perception Remota* (Spanish), *Sensoriamento Remoto* (Portuguese), *Telerilevamento* (Italian), indicates both the acquisition from a distance of qualitative and quantitative information concerning any site and the environment and the combination of methods and techniques for subsequent elaboration and interpretation. If the acquisition occurs from a short distance, the term *proximal sensing* is used.

The combination of measurable parameters from satellites and airplanes is widespread, and applications of remote sensing in environmental sciences are scattered throughout the studies of biology, geochemistry, geology, mineral exploration, geomorphology, lithology, hydrology, oceanography, geobotany, classification of agricultural and forestry resources, environmental pollution, urban planning, risk management, permanent land monitoring, just to mention a few.

Remote sensing does not collect direct information on the environment so the electromagnetic information must be converted into estimates of chemical, biological and physical variables through the creation of appropriate multidisciplinary models. The success or failure of remote sensing in estimating environmental alterations thus depends on models and algorithms developed and used to extract environmental parameters from the *continuum* of collected optical–spectral data from the sensors, and on comparison with Surface Reference Data assessment.

The collection and the distribution of information relate to the development of sensor specifications, data transmission and data processing. Sensors supply remote measures, basically observing the spectral behaviour of the objects in the visible, infrared and microwaves electromagnetic intervals; such measures are used to indirectly adduce the structure of territorial elements or to survey some physical characteristics, like the temperature or spatial distribution of an element. In this regard, far remote sensing permits a qualitative and descriptive analysis of the images, and, under defined conditions, a quantitative, automatic analysis.

Despite the many limitations of remote sensing, a considerable increase has been witnessed both in the number of missions and in the achievable geometric, temporal, radiometric and spectral resolutions, which have enhanced its role in territorial planning, in managing land resources and in monitoring environmental dynamic processes.

Satellite data, concerning the same area, are collected in a period that may vary from a few hours to some weeks. This provides the possibility to deliver an updated thematic cartography, where the only limit is represented by the geometric resolution of sensors with respect to the geometric precision of a map.

1.6 Global Satellite Positioning Systems

The positioning systems applied to points on the Earth's surface have found practical application in topography and cartography, after their initial specific utilization in the field of ship navigation.

They allow the 3D positioning of static or moving objects in space and time, in every place on the globe, under all meteorological conditions and continuously.

They are based on the reception of radio frequency signals emitted from artificial telecommunication satellites. The ground station must be equipped with an antenna and a receiver: their degree of complexity and cost depends on the measurement level of accuracy needed to determine geocentric coordinates (WGS84) of any point on the Earth's surface.

If we know the geocentric position of the satellites, whose orbits are referred to the geocentric reference system WGS84, the geocentric coordinates can be directly converted by the receiver into other reference systems, supplying a 3D positioning.

Defining the position of a point is possible by calculating the distance between satellites and receiver, indirectly determined through measures of time or phase by exploiting the different characteristics of the signal emitted by the satellites and received by the receiver. The nature of such signals is defined from the positioning system of reference.

Two functioning constellations of positioning satellites, planned and launched in orbit during the 1970s and 1980s by the two former political blocs (USA and USSR), currently guarantee position measurements:

- American system NAVSTAR GPS, *NAVigation Satellite Timing And Ranging Global Positioning System*;
- GLONASS, *Global' naya Navigatsionnaya Sputnikovaya Sistema*, currently managed by Russia.

Europe is also working on an alternative constellation for global positioning with the aim of reducing dependence on American and Russian systems and to produce a more modern, reliable and multifunctional system with wide strategic and economic benefits.

On July 19, 1999 the Council of Europe adopted a resolution to get the *Galileo* system underway in cooperation with the European Space Agency (ESA). This system is planned to have complete complementarity with the existing GPS and GLONASS systems.

1.7 Laser Scanning

Among other survey disciplines, the laser scanning technique is particularly significant, as it is characterized by the ability to produce complete information and achieve high precision, and by a considerable level of automation and productivity.

> Starting from a laser source, fixed or in motion, ground based or aerial, through the polar detection of a very large number of points surrounding the laser source and the radiometric measure of each of them, it is possible to recreate, nearly continuously, the three-dimensional image of the object or the surface of interest.

The laser scanning techniques therefore represent a meaningful evolution of some aspects of photogrammetry, directly supplying a 3D surface model, traditionally obtained from the stereoscopic elaboration of bi-dimensional images, thereby reducing the involvement of expert interpreters and approaching the total automation of the process.

Technology and computer science aspects characterize this technique, which offers multiple applications in ground and aerial surveys. Its weakness is represented by the complex and at present ill-defined filtering operations necessary to reduce and to select the enormous amount of data collected by the laser system and necessary to recreate the Digital Surface Models (DSMs).

1.8 Geographical Information Systems

As software for the management and manipulation of geographic data, the Geographical Information System (GIS) appeared in the computer market in the mid-1960s. Progress in the field of GIS software has been rather slow in comparison with the information systems used for commercial or financial data. The delayed introduction of the GIS in the market can be explained by the fact that the data and information to be stored in such systems are more complex and difficult to process compared to other types of non-Geographical Information Systems. Geospatial information, or geoinformation, in fact concerns phenomena referenced in planimetry and in altimetry, which are strongly inter-related. The widespread rapid diffusion of GIS starting in the 1990s has enormously increased retrieval, elaboration and analysis capabilities of available data stored in the archives of public administrations, agencies and research institutes that are fundamental for studying and planning the real world.

A GIS (in French: *Sistème d'Information Géographique*, SIG, in German: *Geographisches Informationssystem*, in Italian: *Sistema Informativo Geografico*) can be defined by focusing on the tool-base (Burrough), the organizational (Cowen) or the spatial database (Aronoff) aspects:

- Burrough (1986) first defined the GIS as *a powerful set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world*.
- Cowen (1988) defined a GIS as *a Decision Support System involving the integration of referenced spatial data in a problem solving environment*.
- Aronoff (1989) considered the GIS as *any manual- or computer-based set of procedures, used to store and manipulate geographically referenced data*.
- Burrough (1997), again, defined the GIS as *a powerful set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world for a particular set of purposes*.

The tool-base definition considers that geographic (or spatial) data represent phenomena from the real world in terms of

- position with respect to a known coordinate system;
- attributes when unrelated to position (name, age, height, etc.);
- spatial interrelations with each other that describe how they are linked together (topology describing space and spatial properties) (Burrough, 1997).

1.9 Decision Support Systems and Expert Systems

Cowen's statement leads to a powerful evolution of the GIS, developing in Decision Support Systems (DSS).

A synthesis of Burrough's and Cowen's statements can successfully summarize up the evolution of the Geographical Information Systems:

> A Decision Support System (DSS) is a powerful set of tools for collecting, storing, retrieving at will, processing, transforming and displaying georeferenced spatial data in adequate scenarios of the real world so as to supply the decision makers with objective elements of evaluation of environmental problems.

By using DSS one must be in a position to prevent and foresee territorial and environmental phenomena and to explore several scenarios to obtain an overview of their possible consequences. For example, the ability to foresee when a volcanic eruption or a flood event may happen, its intensity and the extension of the area involved, may be useful in the definition of a plan for evacuating the population.

1.10 Spatial Information

With the constitution of the *Global Spatial Data Infrastructure* (GSDI) Association, formally founded in July 2003, the term Spatial Data Information was officially introduced as information related to the terrestrial Globe in 3D space.

The purpose of the GSDI organization is to promote international cooperation and collaboration in support of local, national and international spatial data infrastructure developments that will allow nations to better address social, economic and environmental issues of pressing importance.

However, some terminological nuances still remain. The common definition *Geographic Information* (GI) is used to indicate everything concerning the 3D positioning and georeferencing of objects on our planet.

In some cases, the terms Territorial Information and Geographic Information are used as synonyms, although their different cultural backgrounds imply a distinction.

Geographic and Territorial cultures are substantially different because of their humanistic or scientific approach, a distinction that still exists in some countries as between Human Geography and Measurement of the Territory.

This difference is also reflected in expressions such as *Territorial Information System* (TIS) and *Geographical Information System* (GIS), although the term GIS is accepted worldwide.

In this volume, reference will often be made to *GeoSpatial Information* (GI) as a good definition of the space we live in, increasingly measured, described and represented in its three dimensions.

1.11 Geography

Geography possesses an intrinsic specificity and an epistemological autonomy such as to offer valid tools for an understanding of the real world.

The study of geography extends to a vast range of natural anthropical phenomena. Geography, in fact, based on the findings of physical and human sciences, performs an interpretative synthesis in order to analyse relationships, causes, effects and evolutionary tendencies.

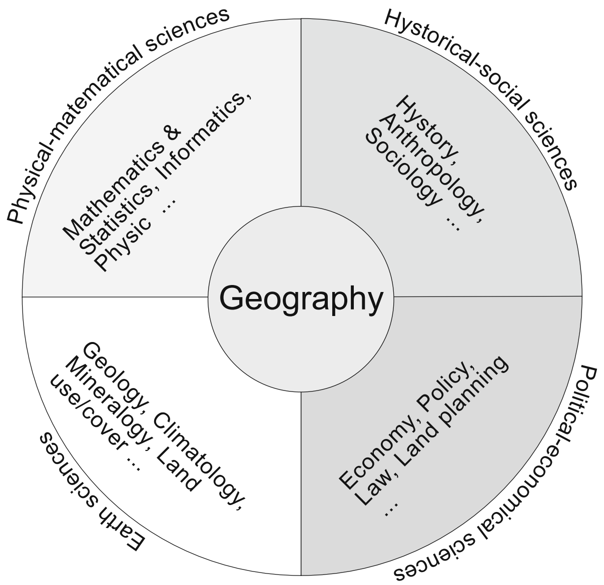


Fig. 1.2 The wheel of geography

Geography is a qualifying and constructive means to achieve correct environmental and territorial policies (Fig. 1.2).

Anthropic geography studies the condition of humans on the territory and the relationships between organized human societies and natural environments, in the planetary geo-system and in each ecosystem at different scales. Perceptions, interactions, hierarchies and different developments correspond to each of these; the location of whatever phenomenon cannot therefore be confined to an absolute analysis; rather, it must be seen in relation to increasing radius situations.

Such a location must not be interpreted as the position of a place only; rather, it must encompass the concepts of distribution, association and spatial specialization. In this way, distance can be not only metrical but also temporal, economical and social.

1.12 Ontology

A brief section is devoted to Ontology in this volume and in this chapter because of the growing interest it is attracting especially within the spatial information domain and because of its role in contributing to ordering terminologies and definitions often inadequately used as a consequence of the rapid development of geomatics.

The term Ontology is often controversial in discussions of Artificial Intelligence (AI). It dates back to the early stages of philosophy, with particular reference to the view of science as a historical fact and as an expression of human rationality. Ontology is very often confused with Epistemology and it is therefore worth explaining the philosophical meaning of the two terms:

- *Epistemology:* the study of nature and of the value of scientific knowledge, sustained by critical examination;
- *Ontology:* also known as metaphysics, i.e. a philosophical doctrine related to the universal characters of the Supreme Being (God), corresponding to the older Aristotle's first philosophy. Ontology is the specification of conceptuality and the description of concepts and the relationships existing for an element or between elements of a group, entity or class.

The common denominator, of those who ventured into these thought traditions, is the account of theories of knowledge dealing with the history of scientific knowledge and with the development and consequent impact on society and the planet.

What is the role of ontology in geomatics? Ontology in geomatics implies the definition of different ontologies involved with sharing and re-using knowledge, so as to help define ontological assumptions.

An ontological assumption is the agreement to use a dictionary (asking questions and making statements) in a consistent but incomplete way, in comparison with the theory specified by ontology. Some elements reflect preset ontological principles. Sharing knowledge with and among these elements is possible thanks to the implemented ontologies.

The formal representation of knowledge is based on conceptualization: on the one hand there are objects, concepts and other entities assumed to exist in a certain domain of interest, while on the other hand we have the relationships established among them. The conceptualization is an abstract, simplified vision of the world that we wish to represent for a certain purpose. Every basic element of knowledge, every system based on knowledge and every expert system is driven by an explicit or implicit conceptualization.

In artificial intelligence, what exists can also be represented. When knowledge within a certain domain is represented in a descriptive formalism, the group of objects to be represented is defined as universe of discourse. A dictionary is used to describe the relationships among objects belonging to the same group, which is also used to represent knowledge itself by means of a program based on knowledge.

In the context of artificial intelligence, it is possible to describe the ontology of a program by defining a group of representative terms for that application. In such an ontology, definitions associate the names of entities in the universe of discourse (for example, classes, relationships, functions, other objects) with a descriptive text showing the exact meaning of the terms, assuming formal axioms, or limiting possible interpretations and use of such terms.

From a formal point of view, ontology is the definition of a logical theory.

The simplest example of ontology is the hierarchical taxonomic definition of classes in a legend (i.e. CORINE LandCover).

1.13 The Geomatics Expert

The increasing interest in the field of spatial information, the complementarity, the integration and the synergy among the disciplines and techniques characterizing it have recently led to the characterization of a new professional profile: the *Geomatics Expert*.

A similar professional profile already existed in North America at the beginning of the 1980s, whence it and progressively circulated in Europe and worldwide. It is the evolution of the *Chartered Surveyor*'s profile in the English-speaking countries, of the *Géométre Expert* in the French-speaking countries, of the *Vermessungsingenieur* in the German-speaking countries and of the *Geodetische Ingenieur* in the Netherlands.

A careful reading of the historical itinerary and different evolution processes identifies mathematics, physics and particularly astronomy as the common origin of the survey disciplines, later to be elevated to sciences of land surveying starting in the 17th century.

The first differentiation takes place between the end of the 18th and the beginning of the 19th century in the'France of the Revolution and of the Empire', when the survey disciplines were included in the emerging field of civil engineering.

Only later, did other differentiations occur at the European level.

The development of survey disciplines has progressed rapidly over recent decades: from spatial geodesy to precision topography, from photogrammetry to

remote sensing, from numerical cartography to the processing of observations, to the GIS, DSS and expert systems.

Cartographic institutions have taken upon themselves official responsibility, increasingly extended, for the problem of quality certification (in survey and in projection).

The professional prospects for the expert in surveying and monitoring are therefore growing continuously and progressively. The professional profile of the geomatics expert is rapidly emerging, and this may correspond to a job opportunity, if adequately supported by high-quality education and training.

1.14 Summary

The term ‘Geomatics’ indicates an integrated, multidisciplinary approach to the use of specific tools and techniques chosen to acquire, integrate, manage, analyse and spread georeferenced spatial data in digital format and in a continuous way. The innovation of this term is the fact that any geographic element on the Earth can be georeferenced.

Geographic data, which is a measurement of an object or of one of its characteristics, were converted by expert systems into information, which can be used for several applications.

Geomatics is constituted by many disciplines, which integrate with each other in the detection, collection and representation of a big amount of data to produce different kinds of information. These disciplines are

Computer science, which studies algorithms that describe and transform the information; it provides both processor technology (hardware) and methodological tools and systems to generate the information (software);

Geodesy: the science that deals with shape and dimension of the Earth, in relation to the external gravitational field, and defines the reference surfaces (the geoid and the ellipsoid).

Cartography, which provides plane representations of the 3D surface of the Earth, describing its shape, size and details, according to defined standardized rules.

Photogrammetry, which guarantees the geometric correspondence between the real object and its representation based on photographic images.

Remote Sensing, which includes those techniques and instruments (sensors) allowing the distant acquisition of territorial and environmental data, as well as methods and techniques for data processing and interpretation. It is based on electromagnetic radiance properties of detected objects on the Earth’s surface and provides up-to-date data thanks to the short revisiting times of modern sensors.

Global Satellite Positioning Systems based on the reception of radio signals from artificial satellites for communication, they can retrieve the coordinates

of any object on the Earth's surface, providing continuous 3D positioning, even when the object is moving and in any weather condition.

Laser scanning systems allow the reproduction of the 3D image of an object or a surface, through the use of a laser source which emits a very large number of radiation events in the optical frequencies (0.3–5 μm) of the electromagnetic spectrum and of a sensor receiving the radiometric backscattering of the detected object/surface.

Geographical Information Systems (GIS): a category of software able to collect, store, recall, transform and represent georeferenced spatial data; they include powerful tools allowing complex analysis of large amount of georeferenced data.

Decision Support Systems: these are complex GIS developed to draw scenarios by modelling ground truth to be used by decision makers in environmental problem assessments, providing them with a valid tool to prevent, forecast and properly face the effects of natural events.

Spatial Information is a term that refers to information concerning the Earth in 3D space.

Ontology is the description of the concepts and relationships existing for an element or among elements of a group, entity or class; conceptualization is a simplified, abstract vision of the world to be represented for a certain application.

Given now interest in the domain of Spatial Information has recently increased, a new profile, the geomatics expert, has been defined: he/she is the one who can integrate and manage all the different techniques and disciplines of which geomatics is composed.

Further Reading

Aronoff S., 1989, Geographic Information Systems: A Management Perspective. WDL Publications, Ottawa, Canada.

Burrough P.A., 2000, Principles of Geographical Information Systems, Spatial Information Systems and Geostatistics, Claredon Press, Oxford, p. 306.

Hofmann-Wellenhof B., Lichtenegger H., Wasle H., 2008, GNSS – Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and More, Springer, Wien New York p. 516, ISBN: 978-3-211-73012-6.

Commission of the European Communities, 2007, INSPIRE Directive, 2007/2/EC of the European Parliament and of the Council on March 14 Establishing an Infrastructure for Spatial Information in the European Community, Official Journal of the European Union, ISSN 1725-2555, L 108, Vol. 50, 25 April 2007.

Gomarasca, M.A., 2004, Elementi di Geomatica, Associazione Italiana di Telerilevamento, p. 618.

Gomarasca, M.A., (ed.) 2008, GeoInformation in Europe, Proceedings of the 27th EARSeL Symposium, Millpress, The Netherlands, p. 681, ISBN 9789059660618.

Kraus K., 1993, Photogrammetry. Vol. 1. Fundamentals and Standard Processes. Dummler, Bonn, p. 397.

Mikhail E.M., Bethel J.S., McGlone J.C., 2001, Introduction to Modern Photogrammetry. John Wiley & Sons Inc., New York.

Richards J.A., Xiuping J., 1998, Remote Sensing Digital Image Analysis, An Introduction, 3rd revisited and enlarged edition. Berlin Heidelberg Springer.

Swain P.H., Davis S.M., 1978, Remote Sensing: The Quantitative Approach. Mc Graw-Hill Book Company, Paris, New York.

Bibliography

Annoni A., 2003, INSPIRE: Infrastructure of Spatial Information in Europe, 9th EC GIS&GI A Coruna, Espana, 25–27 June 2003.

Annoni A., 2007, The Contribution of INSPIRE to Sustainable Development, Proceedings of the 27th EARSeL Symposium, GeoInformation in Europe, M.A. Gomarasca (Ed.). Millpress, Netherlands, pp. 33–20, ISBN 9789059660618.

Burrough P.A., 1986, Principles of Geographical Information Systems for Land Resources Assessment. Clarendon Press, Oxford.

Cowen D.J., 1988, GIS versus CAD versus DBMS: What are the differences? Photogrammetric Engineering and Remote Sensing, 54: 1551–1554.

Gomarasca M.A., 2000, Introduzione a Telerilevamento e GIS per la Gestione delle Risorse agricole e Ambientali, Ed. AIT, p. 250 pp, 32 Tavole a colori, 2nd Ed.

Gomarasca M.A., 2007, International Review of Geo-spatial Information and the 7th European Framework Programme, Proceedings of the 27th EARSeL Symposium, GeoInformation in Europe, M.A. Gomarasca (Ed.). Millpress, Netherlands, pp. 3–11, ISBN 9789059660618.

Gruber T.R., 1993, A translation approach to portable ontologies. Knowledge Acquisition, 5(2): 199–220.

Gruber T.R., 1993, Toward principles for the design of ontologies used for knowledge sharing. Presented at the Padua Workshop on Formal Ontology, March. Nicola Guarino Edited collection.

ICA, 1995, 17th International Cartographic Conference, 10th General Assembly of the International Cartographic Association, Barcelona, Spain, 3 September. Springer, The Netherlands, ISSN: 0343-2521.

Michalsky R.S., 1994, Inferential Theory of Learning: Developing Foundations for Multistrategy Learning, Machine Learning – A Multistrategy Approach, Michalski R. and Tecuci G. (Eds.), Vol. 4. Morgan Kaufmann Publishers, San Francisco.

Philipson W.R., 1997, Manual of Photographic Interpretation. American Society for Photogrammetry and Remote Sensing, 2nd Edition, pp. 3–20.

Wolf P.R., 1988, Elements of Photogrammetry, Cap. 3: Principles of Photography. McGraw-Hill International Editions, McGraw-Hill Book Company, International Editions, New York pp. 41–60.

Chapter 2
Elements of Cartography

Cartography is of fundamental importance for all geographic data characterized by a spatial attribute, i.e. datum coordinates, as it provides a possible description of the shape and size of the Earth and its natural and artificial details. The role of cartography is to map geographic facts and phenomena (e.g. places) which are identifiable by their position and which can be georeferenced, archived and referred onto maps.

The basic functions of cartography are

- to give both a *punctual* knowledge of the territory, based on the observation of every single object, and a *general* one, for a global vision;
- to allow and develop deductive and inductive logical processes, related by *concomitance*, *proximity*, *frequency*, etc.;
- to act as basic support for *classification*, *planning* and *land management*.

A cartographic map is a flat, approximate, reduced and symbolic representation of the Earth's surface or of its parts. This is achieved according to systems, or representations, that turn ellipsoidal figures into plane figures and that, through appropriate signs, can reproduce information about the shape and measure of the ground elements, as well as the relative representation of real or abstract phenomena which can be located in space.

A map is a cartographic product realized by integrating different disciplines and techniques. The phases of the cartographic process can be summarized as follows:

- survey, selection and annotation of spatial data;
- data standardization;
- data generalization;
- transformation of the points belonging to the terrestrial surface into the corresponding points on the *reference surface* (geoid: altimetry, ellipsoid: planimetry) through biunivocal correspondence. The terrestrial surface, and of the gravity field, is approximated by choosing a specific coordinate reference system;
- representation of the reference surface on the map according to a *cartographic reference system*, namely the mathematical equations enabling one to represent the terrestrial surface on map plane or screen;
- realization of the map, defining its typology, legend and symbols.

M.A. Gomarasca, *Basics of Geomatics*, DOI 10.1007/978-1-4020-9014-1_2,

Nowadays, the photogrammetric techniques, as well as the most recent remote sensing techniques and laser scanning system data, are replacing the direct survey, thus simplifying and partly modifying the sequence of the phases.

The term cartography is often used with two meanings:

- *with reference to the discipline*;
- *with reference to the elaborations* (maps) resulting from complex cartographic activities.

In this chapter, issues concerning the determination of the size and shape of the Earth, the geodetic reference systems and the cartographic representations are discussed using a descriptive approach. Some elements needed for reading the topographic maps and interpreting the landscapes are also reported.

2.1 Milestones in the History of Cartography

The history of cartography is the intersection of the disciplines that underlie it. An exhaustive revisit is almost impossible, but some of the facts and the figures philosophers, astronomers, mathematicians, geographers, etc. who have left their indelible signs may be mentioned. The purpose of this section is to highlight how the problem of describing and measuring the Earth had always been of great interest among the researchers through different periods and with different approaches that, particularly in the ancient world, resonate still, even with current knowledge.

The Soleto Map is the oldest geographic map ever discovered in Europe. It is a piece of a crate enamelled in black, a little *ostrakon* of 5.9 cm $\times$ 2.9 cm with the incision of the coastal line of the Salentino peninsula, south of Italy, with some Greek toponyms and 11 local toponyms shown with points. The object was discovered by the archaeologist Thierry van Compernolle of the University of Montpellier in Belgium, August 21, 2003 in a large messapic building, confirming the relationship among Iapigi, Messapi and Greeks in the V century BCE (Fig. 2.1a).

The great prehistoric and historical civilizations of the Near and Middle East offer indications of the existence of a sort of rudimentary cartography. Among the most ancient, in Northern Mesopotamia, a *graffito* has been found, drawn on a clay tablet and representing the Euphrates and a tributary river, the mountains bordering Northern Mesopotamia and the cities, represented by a small circle (Fig. 2.1b and c).

In Valcamonica (Northern Italy), once inhabited by the people of the ancient Camunis, over 170,000 rocks' figures have been dated back to the period between the Palaeolithic and the Iron Age. Among these *graffiti*, in a complex chronological definition, are some *petroglyphs*, which are considered as maps, and which probably describe the articulated disposition of a settlement with houses and workshops, as a highly detailed projection (Fig. 2.2).

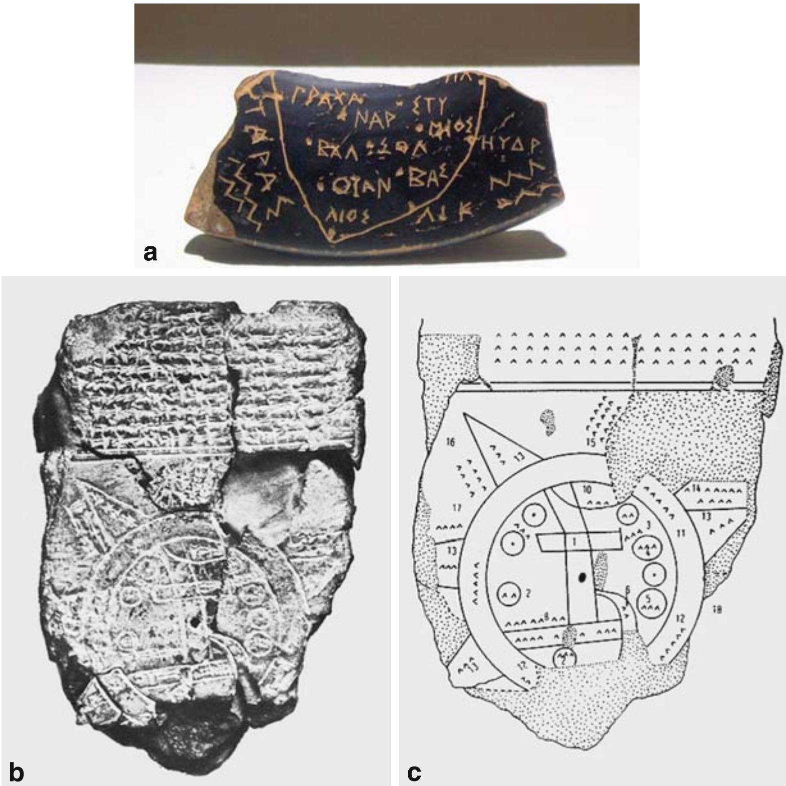


Fig. 2.1 (**a**) The Soleto Map is the oldest geographic map ever discovered in Europe; (**b**) World map inscribed by the Babylonians on a clay tablet representing Mesopotamia (? 2400–2200 BCE) stored in the British Museum in London. This map represents the world as a flat dish surrounded by water, as known at that time (**c**)

The Egyptians even drew cadastral plans for the recognition of property boundaries for cultivated lands as well as a topographic map of a gold deposit in Nubia, preserved in the Egyptian Museum in Turin, and dating back to 1200 BCE.

During the Greek and Roman times, many scholars ventured upon the study of our planet, formulated different hypotheses and proposed solutions for its representation.

Cartography conceived for cultural purposes was first born in ancient Greece. The first map of the world, representing the *ecumene*, namely the emerged land inhabited by man, is attributed to a philosopher of the ionic school, developed along the Aegean coast of Asia Minor, Anaximander of Miletus. He was a disciple of the philosopher and mathematician Thales (624–546 BCE) and drew the map in the mid-VI century BCE (Fig. 2.3), followed by Hecateus's work in the V century BCE.

Cartography began its scientific progress in the IV century BCE, still in classical Greece, where Dicearcus of Messina (IV century BCE), a philosopher belonging to the generation after Aristotle, for the first time introduced a mathematical construct

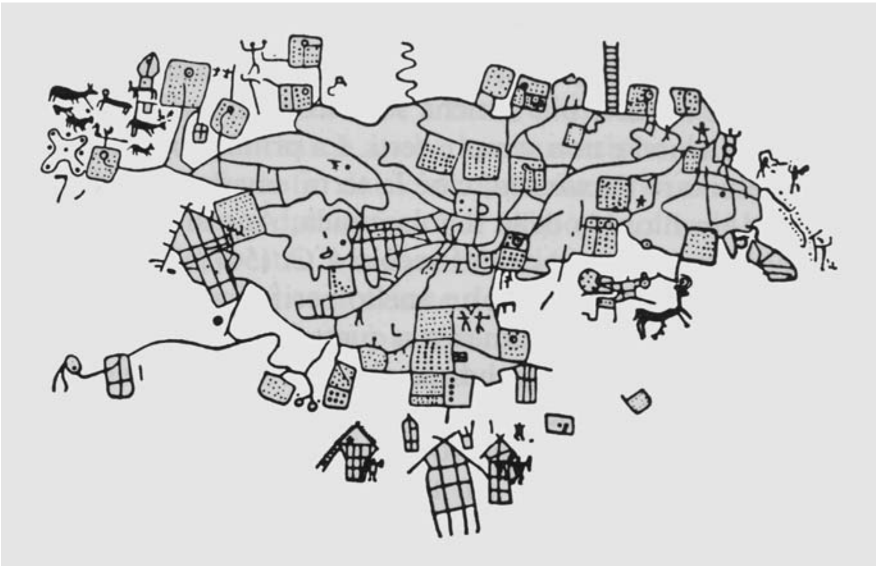


Fig. 2.2 Reproduction of the Camunic (from Valcamonica Valley, Italy) petroglyphs (rock incision): prehistoric map of Bedolina, a village in Valcamonica Valley in the Italian Alps

Fig. 2.3 The Earth according to the representation by Anaximander of Miletus in the VI century BCE

element, constituted by an East–West line, that is a parallel, passing through the Pillars of Hercules, Sicily, Athens, Rhodes, Taurus and Mount Imaus, all places that were considered to lie on the same latitude. Thus was the concept of a reference system introduced: the *diaphragm* having the meaning of parallel and the *perpendicular* acting as a meridian, while the origin of the axes was in Rhodes.

Eratosthenes of Cyrene (III century BCE) around 250 BCE succeeded in calculating the circumference of the Earth, to a good approximation, starting from the observation of the angular difference between the Sun's rays over two points of the terrestrial surface. Among the studies and theories proposed in the past, this is undoubtedly the most fascinating method, both for its clarity of reasoning and for its astonishing precision, despite the limited scientific knowledge of that period.

Eratosthenes calculation based on these known facts:

- at noon, during the summer solstice, the Sun was exactly on the vertical above the town of Siene in Egypt, near present-day Aswan. Eratosthenes had reached this conclusion by observing that the Sun was reflected onto the bottom of a well in the city. In fact, Siene was actually in correspondence to what is nowadays defined the Tropic of Cancer;
- on the same day at the same time, in Alexandria (Egypt), about 5000 stadia (1 stadium: $\sim$185 m) North of Siene, as calculated by Eratosthenes himself, a vertical pole fixed into level ground, projected a shadow whose size indicated a Sun angle equal to $7°12'$.

The deduction was that if the two lines passing respectively across the centre of the well in Siene (indicated with S in Fig. 2.4) and the prolongation of the fixed pole in Alexandria (A in Fig. 2.4) met at the centre of the Earth, the distance between the

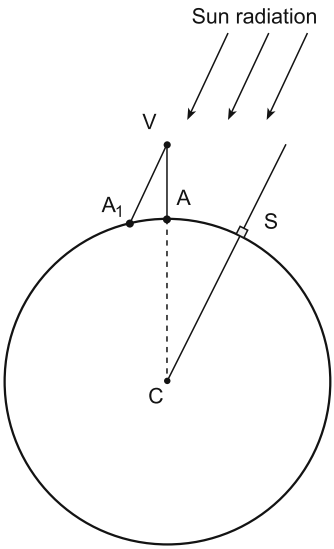


Fig. 2.4 The measure of terrestrial circumference as calculated by Eratosthenes (III century BCE)

two towns was 1/50 of the circumference (360°: 50 = 7°12′): that is 250,000 stadia (5000 × 50) corresponding to 46,250 km, just 15% more than what is measured nowadays.

While it is possible to endlessly discuss the numbers, the units of measurement, the conversions and the deduced results, there is nothing to object to about the method, which is amazingly ingenious.

Eratosthenes also left a geographic description of the world, divided into regions and showing a wider geographic knowledge along the East–West direction, which improved the construction system started by Dicearcus: it has different reference lines at irregular distances, coinciding with the parallels which pass through familiar places, and lines perpendicular to the previous ones, at uneven distances as well, which get closer to the concept of a geographic grid (Fig. 2.5).

Later, Hipparchus of Nicaea (II century BCE) solved the problem of the coordinates of a point on the Earth by applying astronomical methods. He drew up a list of more than 800 stars assigned to six classes of apparent size and measured the distance between the Earth and the Moon, obtaining a result very close to reality. A supporter of the geocentric theory, he constructed the basis of the Ptolemaic system thanks to his studies in geography and cartography; he introduced the use of geographic coordinates and the method of stereographic representation (Fig. 2.6).

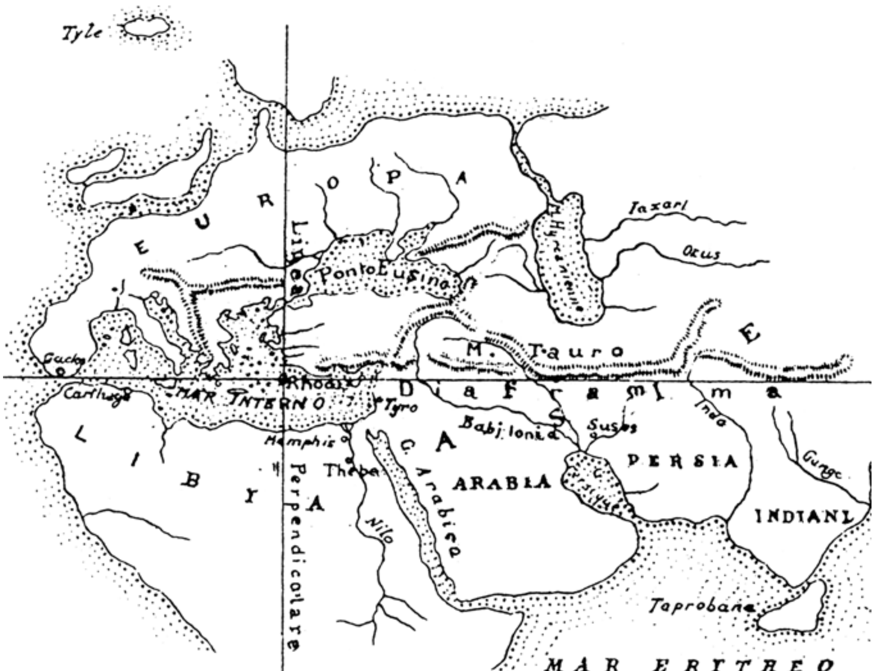


Fig. 2.5 Dicearcus's map (300 BCE) shows the first reference grid system with axes

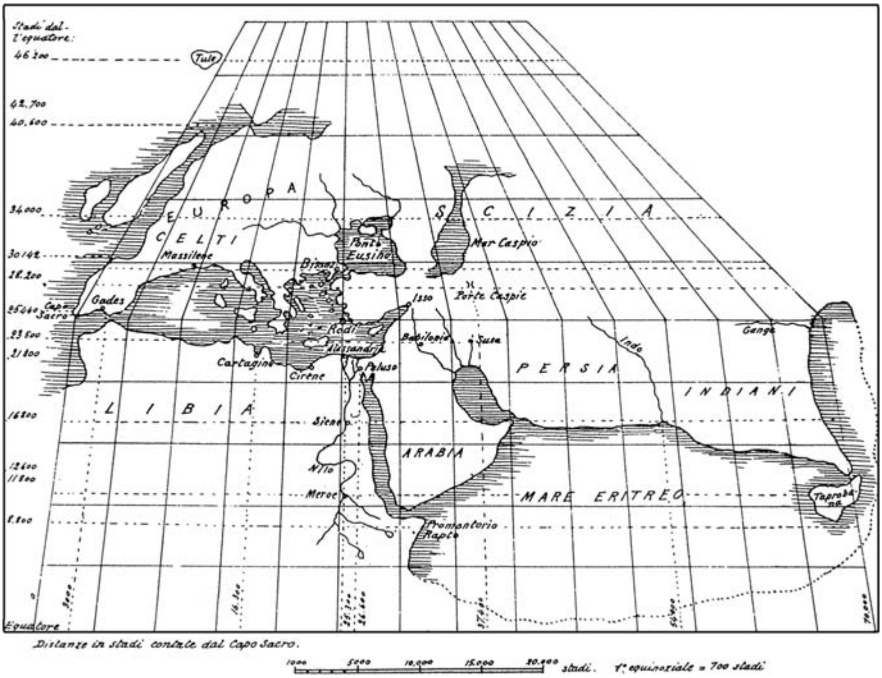


Fig. 2.6 Hipparchus's map (BCE 150). First map with geometric representation and geographical equidistant grid (measured in stadia, 1 stadium: about 185 m)

His works, almost all now lost, were handed down through the papers of Claudius Ptolemaeus (Greek: Κλαύδιος Πτολεμαῖος *Klaúdios Ptolemaĩos*; 83–161 CE), known in English as Ptolemy.

The Roman conquests and the development of trade contributed to the expansion of geographic knowledge, but virtually none of the maps drawn on parchment or papyrus during that period have survived until today. Nevertheless, some traces are left in the second half of the 1st century of the work of land surveyors (*agrimensores*, in Latin), who used to measure the fields and divide into centuries the land subjected to Rome along the main roads, which was the origin to the first wide survey of the territory.

The Romans added nothing to the theoretical basis about the shape and size of the Earth obtained from the Greeks, nor about the distribution of land and water and the ways to represent a spherical surface on a plane. They took credit, in the field of cartography, for their great practical skills, as witnessed by a famous relic dating back to the III or IV century CE, which is known to us thanks to a late medieval copy produced by the geographer Castorius in the XIII century: the *Tabula Peutingeriana* (*Tabula*: geographic map in the Roman period, *Peutingeriana*: from Mr. Konrad Peutinger who found it). It is constituted by a strip, 6.80 m long and 34 cm wide, which represents the outlines of the Roman Empire, from the Iberian Peninsula to the Caspian Sea, including information on roads, towns, distances between them, rivers, mountains, etc.

In the meantime, Marinus of Tyro (I century CE), following the intuitions of the astronomer Hypparcus, used for the first time a grid of reference lines, i.e. meridians and parallels, rectilinear and equidistant. Thus he paved the way to the realization of maps based on a geographic grid, which fixes the different points on the map in relative positions similar to the real ones.

This geometric aspect is particularly evident in Ptolemy's planisphere, the only one from the ancient world, which is based on mathematical and geometric methods for the Earth's representation, namely on a long series of longitude and latitude data, reported on a conical projection grid with circular parallels and meridians converging to the Poles (Fig. 2.7 and Plate 2.1).

The subsequent barbarian invasions caused the loss of a large part of the knowledge acquired in ancient times. Only some notes were saved thanks to the preservation by the Arabs and by Christian monks in the period from V to the XI century.

A renewed interest in cartography was registered in Sicily in the XII century thanks to Abu Abd Mhammad, an Arab better known as Edrisi or al-Idrisi. Thanks to him, geography started to rely on direct and precise knowledge, whereas until that time the influence of imaginary concepts and religious beliefs was of the greatest importance. The planisphere that Edrisi drew bears witness to this fact.

The invention of the compass gave a new impulse to cartography, with Genoese and Venetians travelling to Africa, while the Spanish and Portuguese reached as far as Sudan.

Another important figure in the history of cartography was *Muhiddin Piri Ibn Haji Mehmet* (or Memmed), born in Oelibolu (Gallipoli, Southern Italy) between 1465 and 1470, who died in El Cairo, Egypt, in 1554. He used to be a Turkish pirate, probably of Greek origin. He became Admiral (*Re'is* in Arabic) of the Crescent

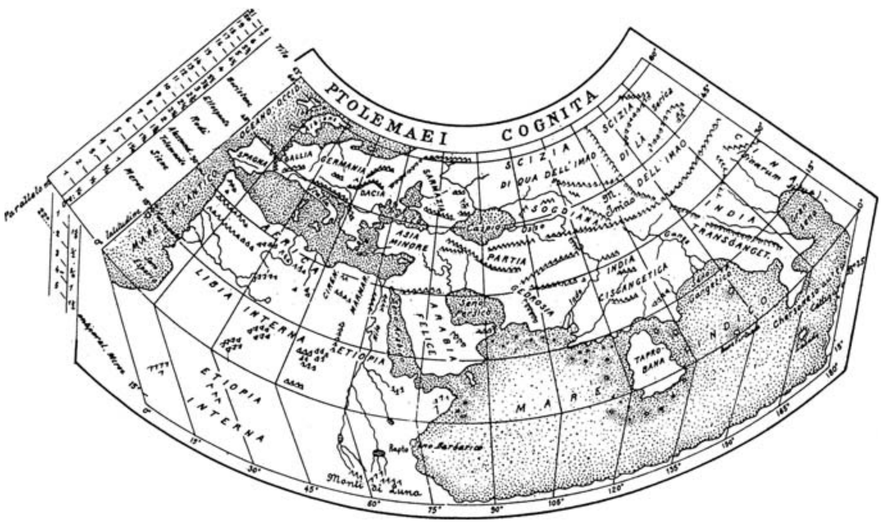


Fig. 2.7 World map in Ptolemy's configuration (CE 150): conic representation (projection) and graduation of meridians (upper border 5° by 5°) and parallels (right border 5° by 5°)

fleet and lived in the time of Suleiman II the Magnificent (1520–1566). As Admiral, Piri Reis had free access to the Imperial Library of Constantinople, where many old maps, never afterwards found, were preserved. Finally, he managed to draw a world map. It was the month of the blessed Muharrem, that is the Muslim year 919, corresponding to the period of the Gregorian calendar which goes from March 9th to April 7th 1513 CE, and the map was given as a gift by the admiral to sultan Suleiman I the Cruel (1512–1520) in 1517. This map, considered one of the first world maps showing the Americas, if not the very first, is surely the most precise map drawn up in the XVI century (see Plate 2.2).

He had a passionate interest for his collection of old maps and was considered an experienced man concerning the lands and coasts of the Mediterranean Sea, so in 1523 he was commissioned by Suleiman to draw an atlas which has become a milestone in the history of modern cartography. Some parts of this atlas are still preserved in the Berlin Museum. This book, intended as a navigation atlas, was the *Kitabi Bahriye* (The Book of the Sea), in which Piri Reis described all the details of coastlines, beaches, currents, bays, straits and shallow waters of the Mediterranean and Aegean Seas.

Reis, constantly searching for new information, also obtained some maps from a mariner who sailed with Christopher Columbus, whose drawings on the parchment, even though yellowed, were very precise.

Be that as he may, the real great innovation in cartography begins in 1569, thanks to Mercator, who first adopted scientific and mathematical procedures to apply a cylindrical, isogonic, increasing latitude projection still known as the conformal cylindrical representation of Mercator.

The very first atlas, intended as a collection of geographic maps, was created in 1570 by Ortelio with the cooperation of most geographers of that period, many Italian among them, who drew the maps according to the knowledge and results of the latest explorations.

More generally, during the Renaissance, Ptolemy's studies were taken into consideration again, and the grid of parallels and meridians became the reference system that is still used today. Between the XVI and XVII century, directly under the influence of Ptolemy's innovation, different kinds of projection were introduced, supported by the development of studies in geometry and pictorial art.

Abraham Ortelio in 1570 arranged the *Theatrum Orbis Terrarum*, the first world atlas.

Gerard Kremer in 1595 introduced the cylindrical projection which still bears his name.

Snellius (XVII century) introduced the trigonometric techniques for the detection of points.

Thereafter, from the XVIII century on, more detailed cartography overcame the restrictions of surveying large properties or military buildings, and land registries started to be compiled.

With the development of new optical instruments, triangulation procedures were introduced, together with the measure of the length of one longitude degree at different latitudes. This was how the realization of modern maps started: in these

maps, most of the main points were determined geometrically. An example is the map of France at 1:86 400, drawn by Cassini between 1744 and 1815.

Lambert (XVIII century) and Gauss later generalized the use of modified representations and originated to national cartographies at small and medium scale (1:100,000–1:25,000). The most commonly used method of tacheometry is a complete survey method scheduling the simultaneous planimetric and altimetric survey of the ground; it is based on the use of relatively advanced topographic instruments.

With the birth of aviation, traditional field surveys were supported by the intensive use of aerial surveys.

With the introduction of the analogical stereo-plotter and the improvement of photogrammetric and image-interpretation techniques in the years between 1920 and 1940, the production of maps started to be definitely based on these instruments and methodologies.

The rest, i.e. the introduction of analytical and digital plotters, electronic processors, aerial and satellite remote sensing, digital sensors, information management and processing capacities, is contemporary history.

2.2 Earth Shape: Ellipsoid and Geoid

The size and the shape of the Earth are topics treated by geodesy. Precision measures, time analysis, distances, stars' position, the force of gravity, etc. allow basic data to be provided for the realization of maps. The principles employed in Earth cartography are also valid and applicable to the cartography of the Solar System (Moon and Planets): Mercury, Venus, Mars, Jupiter, Saturn, Uranus, Neptune and Pluto.

Mankind has always tried to discover the Earth's shape, using different theories. The Babylonians used to think that the Earth was a flat disk, floating on the sea. Aristotle assumed that it could be a perfect sphere floating in space.

Today we know that the Earth has an almost spherical shape which is very regular (Plate 2.3); in fact the distance between the deepest oceanic depth (Mariana Trench: $\sim$ 11,000 m) and the highest peak (Mount Everest: 8,847 m) is equivalent to about 1/322 of the distance between a point at the sea level and the centre of the Earth. A globe with 1 m radius would have, in scale, variations of about 1 mm for the two extreme situations in depth and height: as almost imperceptible depression and relief.

Even hypothesizing that the Earth is completely *smooth*, every point on the terrestrial surface is not equidistant from the core, and gravity's attraction is different at every point due to

- terrestrial rotation around its axis, which generates a centrifugal force (normal to its rotation axis);
- some bulges and flattening respectively, at the equator and at the poles;
- different composition and density of the upper layers.

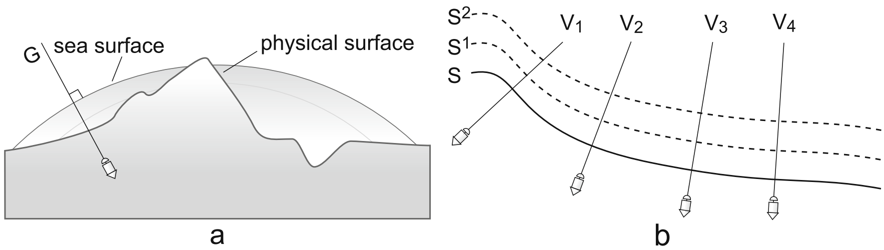


Fig. 2.8 (**a**) Definition of the direction perpendicular to the geoid (positioning a plumbline), (**b**) equipotential surfaces; the geoid is the equipotential surface corresponding to the mean sea level

Correct knowledge of the Earth's gravitational field (treated by geodesy and in particular by gravimetry) unequivocally defines the reference surface of the Earth description: the *geoid* (Fig. 2.8a).

The force lines of the gravitational field (verticals), which are bumpy lines whose direction changes from point to point, together define the field itself and allow us to describe it through equipotential surfaces, or level surfaces (Fig. 2.8b), which are perpendicular to the vertical line at every point. The determination of the vertical at every point of the terrestrial surface is very easy: in fact, it is sufficient to position a plumbline, which unequivocally gives its direction in that point.

Among all the possible equipotential surfaces, the one that has been chosen as reference surface is defined by the mean sea level: the geoid. The geoid is a continuous, smooth surface which is particularly suitable for use as reference. It is found with good approximation based on the surface of the oceans, calculated using points of observations by tide-gauges.

As no reference of the sea and ocean surface exists, measuring the emerged part of the globe requires the use of gravimeters, instruments which can measure the gravity field.

The altimetric position of points lying on the lithosphere is referred to the geoid as an equipotential surface having physical more than geometric consistency.

Hence the geoid is an irregular, theoretical surface, which responds to physical, not mathematical considerations.

The Earth needs to be represented by a shape which allows the positioning of each point through mathematical processes. For this reason, the ellipsoid of rotation, which is the geometric surface that better approximates the geoid, has been defined as the reference surface of the Earth. The idea of spherical geometry was abandoned as the terrestrial surface is characterized by polar flattening, evidence of which has been demonstrated in different historical periods.

As early as 1670, Newton set out a theory to demonstrate that the Earth is not a perfect sphere.

The origin of this geometric anomaly with respect to the spherical model is generally explained in terms of the centrifugal force, C, produced by the rotation of the Earth around its own axis. As it rotates with a constant angular velocity, the

zones closer to the equator tend to move with a higher tangential velocity than the zones closer to the poles. This determines the higher centrifugal force at the equator, which is the cause of equatorial bulges. At the poles, where the tangential velocity is lower, there is stronger gravitational attraction towards the centre of the Earth, which is responsible for the typical polar flattening.

Measurements made by the French Royal Academy of Science in 1735–1743 verified that, with same angular width, an arc of meridian at northern latitudes is much longer than the one at the equator (Fig. 2.9).

The Earth, subject to these forces, has a shape which is very similar to an ellipsoid, which is a geometric surface generated by the rotation of an ellipse around its own minor axis (Fig. 2.10). The ellipsoid represents the ideal surface for cartographic applications, as it can be defined in mathematical terms.

The ellipsoid shape is described by the relative length of the major semiaxis (equatorial) a and of the minor semiaxis (polar) b. The flattening ratio f tends to zero as the ellipse becomes more similar to a circle.

The ellipsoid parameters are

– major semiaxis (equatorial): a
– minor semiaxis (polar): b
– first eccentricity: $e^2 = (a^2\text{-}b^2)/a^2$
– flattening: $f = (a\text{-}b)/a$

Note: in the oldest text books, flattening is indicated by α instead of f.

The ellipsoid equation referred to its own axes in geocentric coordinates is formulated as follows:

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2} = 1 \tag{2.1}$$

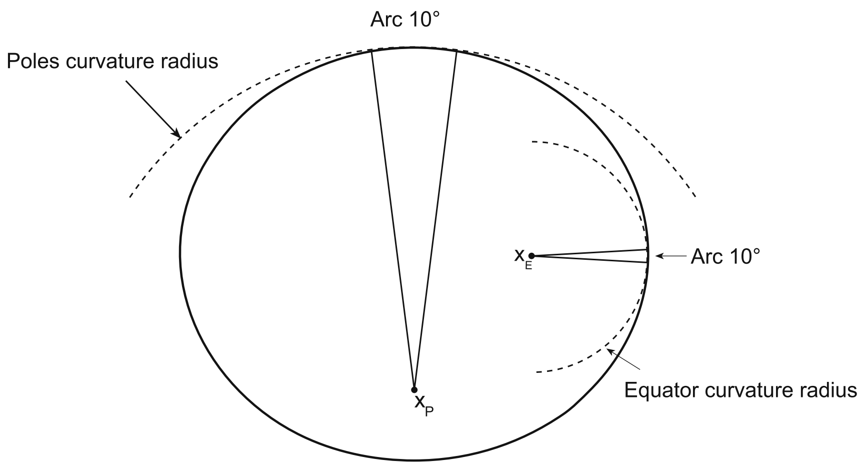


Fig. 2.9 Polar flattening of the Earth's surface. The meridian arc, corresponding to a 10° angle, defines an increasing curve radius to the poles (curvatures exaggerated)

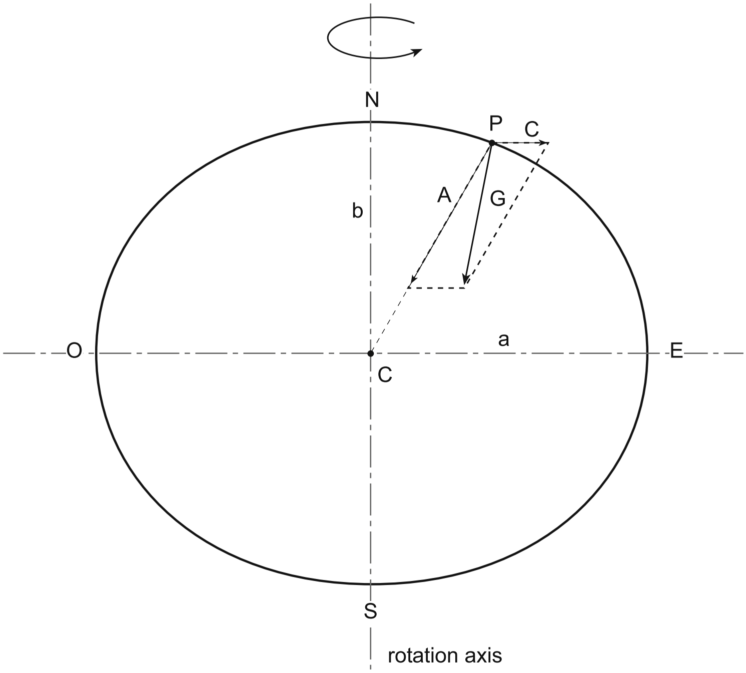


Fig. 2.10 Resultant of the centrifugal force, C, and of attraction, A, at any point of the ellipsoid

The ellipsoid calculation has been carried out through highly precise geodetic calculations, which in general terms follow Eratosthenes' method, namely combining astronomical measurements and terrestrial distances measurements, now perfected.

Over the centuries, different ellipsoids have been described, whose geometric characteristics (Table 2.1) were increasingly appropriate to represent an analytical approximation of the geoid.

The task of geodesy has always been the definition of the rotation ellipsoidal parameters according to various methods used in different periods:

- geometric operations, characterizing meridian and parallel arc measurements in the period from the XVII to XIX century.
- gravity measurements, used in the XX century;
- studies exploiting geodetic satellites, which represent the boundary of current methodologies, as GOCE (*Gravity field and steady-state Ocean Circulation Explorer*): the ESA satellite gravimetric mission.

The *Weights and Measures Commission*, in 1799, gave the measure of 6,375,739 m for the ellipsoid equatorial axis. In the following years, different measures were suggested (Everest, Bessel, Clarke, Helmert) until the Hayford measure was

Table 2.1 Common terrestrial ellipsoidal dimensions used in the past and the WGS84, current reference system, adopted internationally. The Earth's lithosphere is slightly flattened along the polar axis

Ellipsoid	Equatorial semiaxis (a) (m)	Reciprocal flattening (1/f)	$\Delta_{WGS84}a$ (m)*	$\Delta_{WGS84}f$ (10^4)**
Airy	6,377,563.396	299.3249646	573.604	0.11960023
Australian National	6,378,160	298.25	−23	−0.00081204
Bessel 1841	6,377,397.155	299.1528128	739.845	0.10037483
Bessel 1841 (Nambia)	6,377,483.865	299.1528128	653.135	0.10037483
Clarke 1866	6,378,206.4	294.9786982	−69.4	−0.37264639
Clarke 1880	6,378,249.145	293.465	−112.145	−0.54750714
Everest	6,377,276.345	300.8017	860.655	0.28361368
Fischer 1960 (Mercury)	6,378,166	298.3	−29	0.00480795
Fischer 1968	6,378,150	298.3	−13..0	0.00480795
GRS 1967	6,378,160	298.2471674	−23	−0.00113048
GRS 1980	6,378,137	298.2572221	0	−0.00000016
Helmert 1906	6,378,200	298.3	−63	0.00480795
Hough	6,378,270	297	−133	−0.14192702
International	6,378,388	297	−251.0.0	−0.14192702
Krassovsky	6,378,245	298.3	−108	0.00480795
Modified Airy	6,377,340.189	299.3249646	796.811	0.11960023
Modified Everest	6,377,304.063	300.8017	832.937	0.28361368
Modified Fischer 1960	6,378,155	298.3	−18	0.00480795
South American 1969	6,378,160	298.25	−23	−0.00081204
WGS 60	6,378,165	298.3	−28	0.00480795
WGS 66	6,378,145	298.25	−8	−0.00081204
WGS-72	6,378,135	298.26	2	0.000312106
WGS-84	6,378,137	298.2572236	0	0

* ΔWGS84a is the WGS-84 equatorial semiaxis minus the specified datum equatorial semiaxis

** ΔWGS84f is the WGS-84 flattening minus the specified datum flattening multiplied by 10^4

internationally adopted in 1924 at the suggestion of the International Geodetic and Geophysics Commission.

The geodetic geocentric ellipsoid WGS84, its last realization dating to 2004, was finally defined through modern measurement systems and is still current.

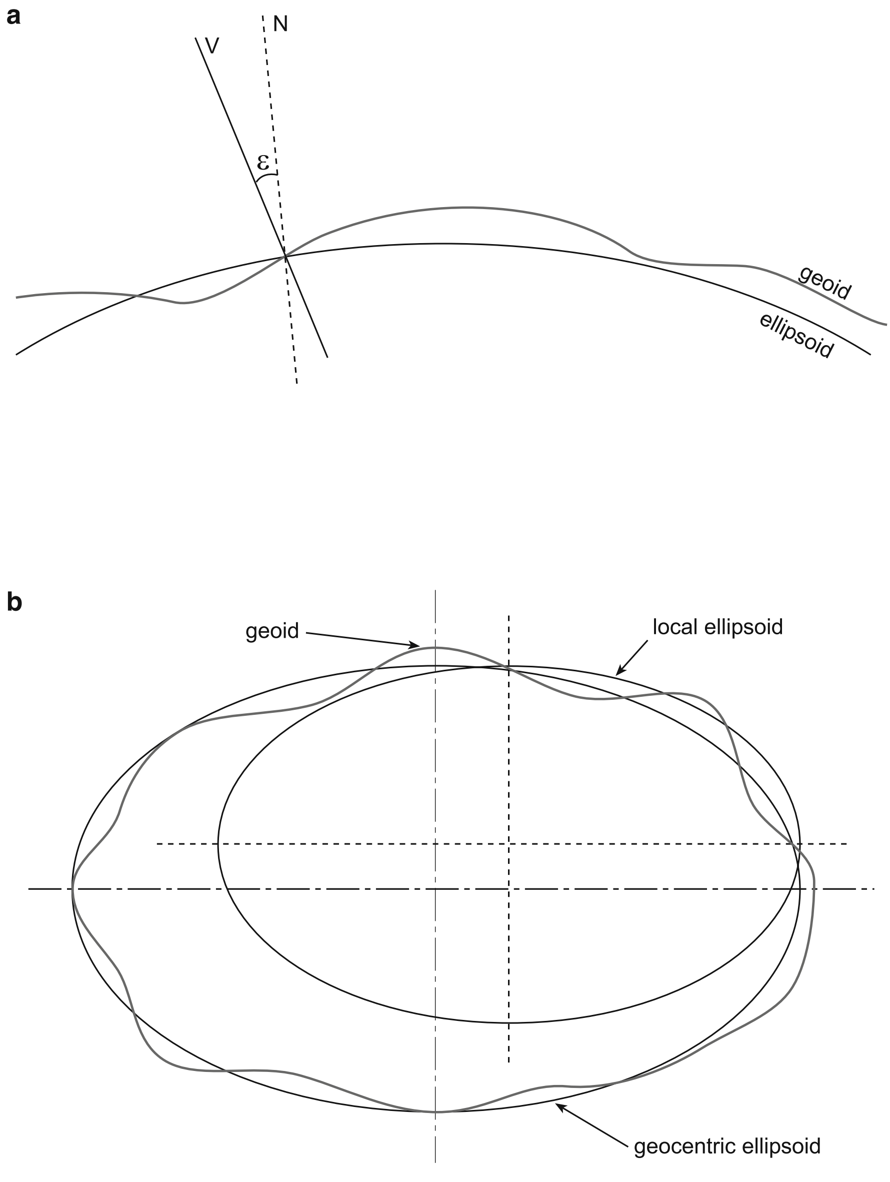


Fig. 2.11 (**a**) Deviation from the vertical at a point belonging both to the geoid and the ellipsoid; (**b**) Geoid, local ellipsoid, geocentric ellipsoid

From an historical point of view, but still with serious repercussions on current national cartographies, the problem of defining reference ellipsoids was traditionally dealt with the *local* scale, separating the planimetric aspect from the altimetric one through the definition of local ellipsoid and geoid. Only the introduction of satellite positioning systems allowed the use of global geocentric systems.

The national cartography of all countries still refers to ellipsoids with a local orientation, so as to better approximate the region represented. It often happens that local reference systems use ellipsoids with common parameters with geocentric ones, but with different orientation; if that is the case, it is called the *datum*.

Whatever is the ellipsoid adopted, the problem of altimetry is still related to the knowledge of the geoid, particular to the evaluation of the deviation from the vertical. The vertical, V, normal to every point of the geoid surface, generally does not coincide with the normal, N, to the ellipsoid surface; the angle formed by the two directions is defined as deviation from the vertical (ϵ) (Fig. 2.11a).

Its determination is one of the goals of geodesy: the knowledge of the deviation of a large number of points uniformly distributed over the terrestrial surface allows one to retrieve the geoid shape in relation to that of a specific ellipsoid. Nevertheless, thanks to sophisticated satellite measurement, an almost direct determination of the geoid is possible. Without these data, it would not be possible to link the orthometric (referred to the geoid) and the ellipsoid heights (Fig. 2.12).

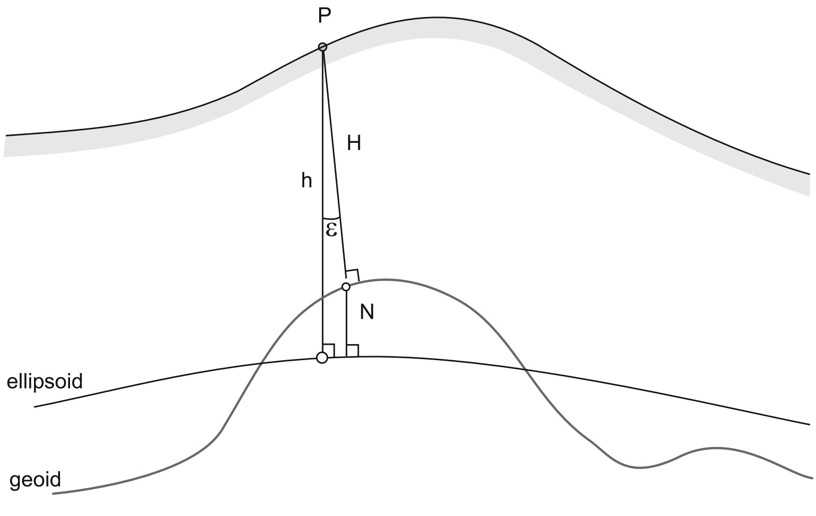


Fig. 2.12 Orthometric height (H), ellipsoidic height (h) of the point P and geoid undulation (N); H = h – N

2.3 Reference Systems

A reference system is the set of measures and rules for the positioning of the terrestrial surface points in space, according to an established coordinate system.

Once the physical and geometric representations (respectively *geoid* and *ellipsoid*) of the terrestrial surface are defined, the problems are the bi-dimensional representation on the map plane and the relations among the different reference systems involved.

The combination *geoid–ellipsoid* is reflected, in terms of points positioning, respectively in the solution of the *altimetric–planimetric* problem.

The determination of a convenient reference system requires one to define a *global/geocentric* reference surface (ellipsoids with geocentric orientation) or a *local* one (ellipsoids with local orientation).

The choice of the surface, the determination of its dimensional parameters and the definition of convenient geocentric, Cartesian, geographic, etc. coordinate systems do not solve the problem of planar cartographic reference systems which, on the contrary, are based on different types of representations.

The problem of the most suitable modelling of the Earth's surface, related to the irregularity of the geoid, determines a non-univocal identification of the reference surface. In a national and international context, this leads to a large number of ellipsoids and their positioning the space so as to locally realize the best approximation to the geoid.

2.4 Ellipsoid and DATUM

Before the high demands for global systems valid for the Earth's surface related to new technologies, like space geodesy and the use of the satellite positioning systems, every nation adopted its own reference system, orienting an ellipsoid with respect to a point defined on their territory, called the *emanation point*.

An ellipsoid is defined as oriented with reference to a point on the Earth's surface when it satisfies the following two conditions:

- it is tangential to the geoid in that point;
- the deviation between the vertical of the geoid and the normal to the ellipsoid assumes constant values at that point.

A planimetric *datum* is the mathematical model of the Earth, defined by a set of rules and measures, used to calculate the geographic coordinates of the points.

The geodetic reference system, *datum*, is constituted by fixing the following elements:

- ellipsoid;
- point of emanation;
- azimuth.

In practice, it is constituted by

- eight parameters: two parameters of ellipsoid shape, six parameters for position and orientation (directing cosines of a semiaxis);
- points compensated network, extended through the area of interest, which fixes it.

In the same *datum*, many coordinate systems can be used: the transformations from one system to another are always entirely mathematical and do not require the introduction of measures. By contrast the transformation between two *datums* can be calculated only when there are enough measures linking points in the two systems.

2.5 Coordinate Systems

The coordinate systems which can describe the position of a point in relation to the reference surface are defined on the chosen *datum*. Knowledge of these systems is fundamental to all procedures of data georeferentiation and for transformation from one reference system to another.

2.6 Ellipsoidic (or Geodetic or Geographic) Coordinates

The geodetic, or geographic, reference system is a particular reference system for the positioning of points on the Earth's surface, based on angular units.

The position of a point in space is determined through latitude and longitude angular measures (Box 2.1) and through the ellipsoidic height (Fig. 2.13); this way, planimetric and altimetric problems are separated.

The determination of the geocentric geodetic systems, like WGS84, is obtained through parameters like:

- angular velocity ω;
- gravitational constant GM;
- second-degree gravitational normalized coefficient C20.

Box 2.1 Geographic definitions

Poles	Intersection of the rotation axis with the ellipsoidal surface
Equatorial plane	Plane that pass through the centre of the ellipsoid perpendicular to the rotation axis
Equator	Intersection of the equatorial plane with the ellipsoidal surface; imaginary great circle around the Earth, everywhere equidistant from the two geographical poles and forming the baseline from which latitude is reckoned. The equator, which measures ~40,076 km, is designated as lat. 0°. The degrees increase to the north and to the south, reaching 90° at the poles
Meridian	Imaginary north–south line between the North Pole and the South Pole that connects all locations with a given longitude (λ). The position on the meridian is given by the latitude. Each is perpendicular to all circles of latitude (ϕ) at the intersection points. Each is also the same size and is half of a great circle on the Earth's surface
Circle of latitude or parallel	Imaginary east–west circle that connects all locations with given latitude (ϕ). The position on the circle of latitude is given by the longitude (λ). Each is perpendicular to all meridians at the intersection points. Those parallels closer to the poles are smaller than those at or near the equator
Longitude λ	Angular distance on the Earth's surface measured along any latitude line such as the equator east or west of the prime meridian (located by international agreement in Greenwich England, on the original site where the meridian passes through the Royal Greenwich Observatory). A meridian of longitude is an imaginary line on the Earth's surface from pole to pole. All points along it are at 0° longitude
Latitude ϕ	Angular distance of any point on the surface of the Earth north or south of the equator. The equator is latitude 0°, and the North Pole and South Pole are latitudes 90°N and 90°S, respectively. The length of one degree of latitude averages about 110 km

As no particular meridian is identified by a specific characteristic, since all are identical, in 1884 the meridian passing across the Observatory of Greenwich in Great Britain was conventionally adopted as starting point 0°, i.e. the fundamental meridian or first meridian. Longitude increases progressively eastwards and westwards up to 180°, completing an angle of 360° in the two directions. In relation to time series measures, the meridian is slightly shifted, as noted in the following observations regarding the terrestrial axis variations.

According to these rules, it is possible to generate a geographical grid made up of meridians and parallels having the following characteristics:

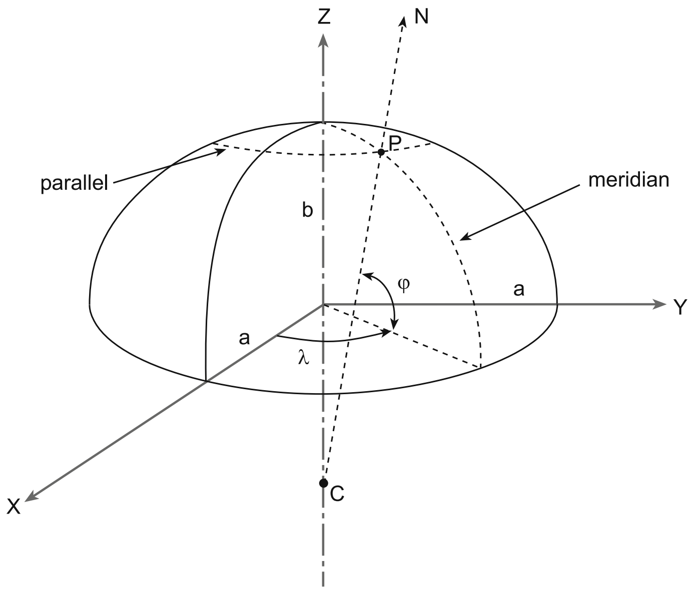


Fig. 2.13 Longitude (λ) and latitude (φ) of a point P on the ellipsoid

- meridians and parallels intersect with a 90° angle;
- number of possible meridians and parallels is infinite.

Meridians are

- half of an ellipse from South Pole to North Pole;
- at maximum distance from one other at the equator and converging at the Poles.

Parallels are

- complete minor circles except for the equator which is a complete maximum circle;
- always parallel to each other (Fig. 2.14).

For a curved surface, like rotation ellipsoids, we have the following reference geographic coordinates:

- longitude (λ): measured in degrees in the range λ: ±180° with respect to the Greenwich fundamental meridian (λ: 0°) or other local ones (Monte Mario, Iron Island, etc.)
- latitude (ϕ): measured in degrees with respect to the equator. It ranges from ϕ: +90° at the North Pole and ϕ: –90° at the South Pole, with ϕ: 0° at the equator.

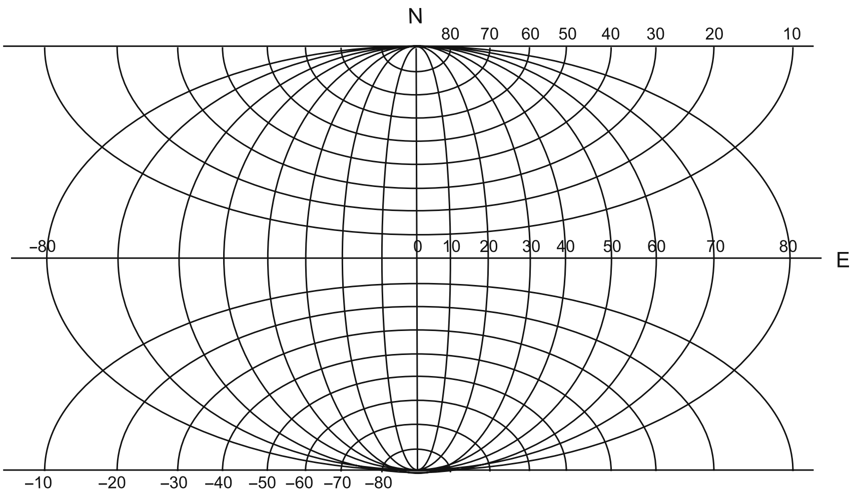


Fig. 2.14 Geographic grid represented by meridians and parallels in the Gauss conformal projection

2.7 Cartesian Geocentric Coordinates

The spatial reference of a point is given by a triple of Cartesian coordinates (X, Y, Z) referred to a system having its origin in the centre of the ellipsoid (Fig. 2.15b).

An example of *Cartesian geocentric coordinates* is the WGS84 reference system (World Geodetic System 1984) in which the Cartesian triple (X_{WGS84} Y_{WGS84} Z_{WGS84}) has the following characteristics:

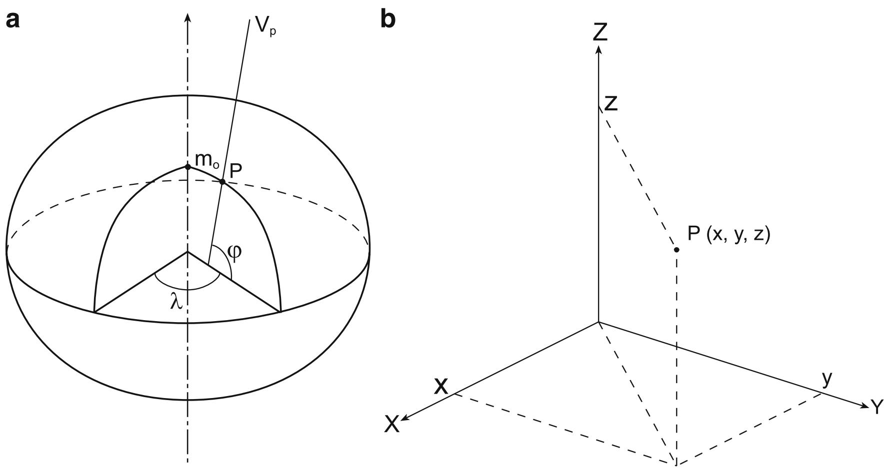


Fig. 2.15 Geographical or ellipsoidal coordinates (**a**) and Cartesian geocentric coordinates (**b**)

- origin at the terrestrial centre of mass;
- Z_{WGS84} axis, oriented towards the position of the medium pole defined in 1984 by the *Bureau Internationale de l'Heure*;
- X_{WGS84} axis, defined in 1984 by the intersection of the medium equatorial plane with the plane of the Greenwich meridian;
- Y_{WGS84} axis, oriented in order to complete a clockwise turn, lies on the equatorial plane.

WGS84 system is used for positioning through GPS (Global Positioning System) technology.

2.8 Planar Cartographic Coordinates

These are the East (E) and North (N) coordinates which determine a reference on the Cartesian plane; they are linked to the geographic coordinates through known analytical transformations whose shape depends on the projection adopted. The planar cartographic coordinate system, for instance in Gauss representation, has

- axis (E): coinciding with the transformed equatorial tract for the projection zone;
- axis (N): coinciding with the transformed zone central meridian.

Mixed coordinates systems are used both in topography and in geodesy: this happens because planimetric and altimetric problems are treated separately, using the geographic coordinates referred to the ellipsoid for planimetry and those referred to the geoid for altimetry.

2.9 Cartographic Projection

For cartographic representation purposes, it is necessary to project the points recognized on the ellipsoid onto the map plane through appropriate analytical transformations.

As it is not possible to project or develop a spherical surface onto a plane without altering the figure size and shape, every planar representation of the Earth necessarily introduces distortion.

It is possible to keep some characteristics unaltered to the detriment of some others. According to the properties which are kept unchanged, it is possible to define different types of planar projections:

- *equivalent*: maps which keep unaltered the ratio between the areas of the represented surfaces. In these maps, equal areas on the map correspond to equal areas in reality.

- *conformal or isogonic*: maps which keep unaltered the angles between two directions in reality and on the map. In isogonic maps, the angle between two directions in reality is the same as in the map;
- *equidistant*: representations along a particular direction, such as the equator or a meridian, along which the scale factor is kept constant;
- *aphylaptic*: maps which present all the alterations described above, but trying to keep them as small as possible in size.

Only the planisphere is a scale representation which can be considered at the same time equidistant, equivalent and conform. For planimetric maps, this is not possible; for example, they may be equivalent but not isogonic. Therefore, it is usually appropriate to choose a specific representation according to the use for which the map is meant.

Figures 2.16 and 2.17 show some distortions caused by the cartographic representations.

The cartographic representations are classified according to the technique used to create the grid of meridians and parallels. A first distinction separates the projections into *pure* and *modified geometric* ones, the former being obtained only through the application of geometric principles, the latter being modified using mathematical functions.

The cartographic projections are obtained by

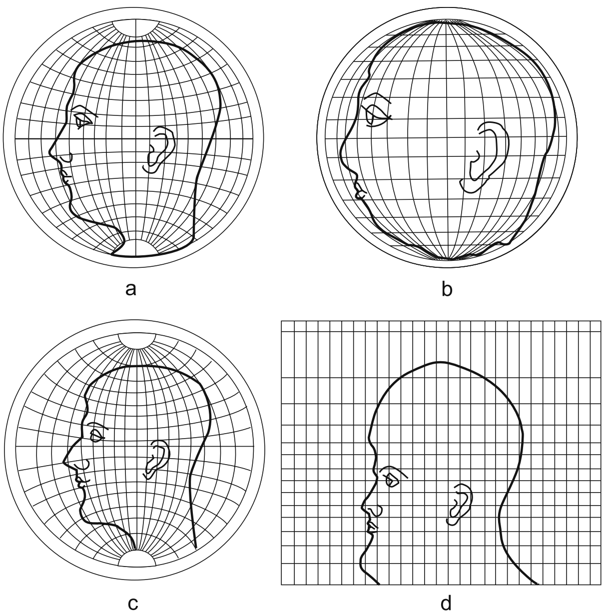


Fig. 2.16 Examples of the distortions (Dents and Adams' head) derived from the projections: (**a**) globular, (**b**) orthographic, (**c**) stereographic, (**d**) Mercator projections

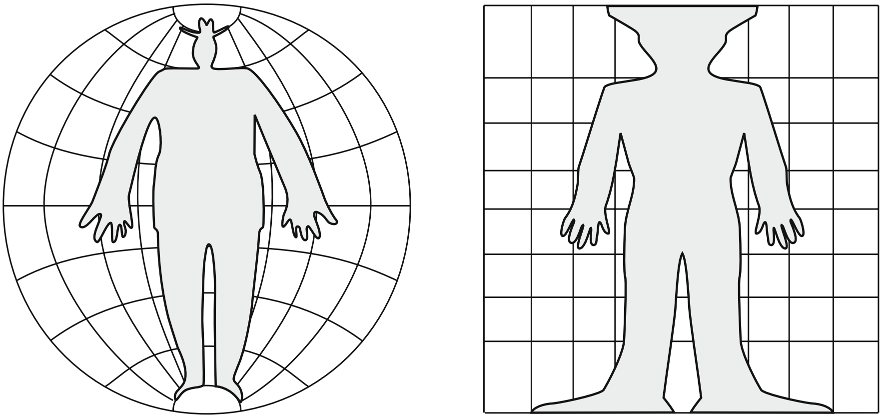


Fig. 2.17 Examples of deformations of a human body introduced by the (**a**) orthographic and (**b**) cylindrical perspective projections

- *perspective projection*, projecting the ellipsoid points onto a plane;
- *development projections* (cylindrical and conical), projecting onto auxiliary geometric surfaces that can be developed into a plane;
- *Pure geometric projections*, e.g. the polar stereographic one, are currently scarcely used.

In general, maps are nowadays generated through analytical projections: Mercator, Gauss (UTM), Lambert representations, developed mathematically.

2.9.1 Perspective Projection

Some perspective projections may be conceived as follows: imagine that the world as a sphere with iron wire meridians and parallels and suppose that a luminous source generates a shadow of this object on a flat screen tangent to the surface itself. This shadow is an example of perspective projection.

According to the position of the luminous source, the projections can be distinguished as (Fig. 2.18):

- *centrographic* or *gnomonic*, with point of view (V) or luminous source in the centre of the Earth (C);
- *stereographic*, with point of view lying on the Earth's surface;
- *scenographic*, with point of view lying outside the Earth's surface;
- *ortographic*, with point of view or luminous source supposed to be at or near infinity and hence with rays or projection lines in parallels.

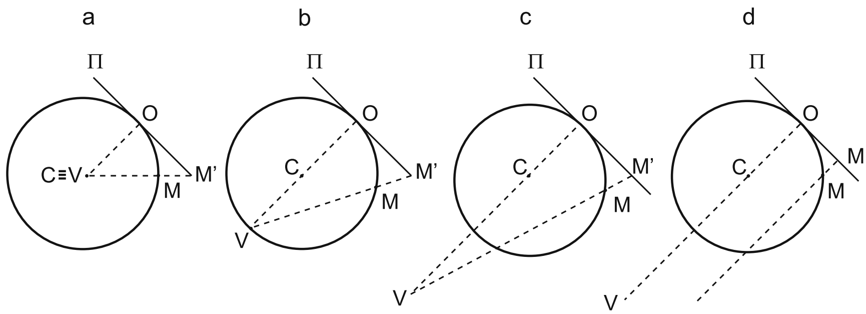


Fig. 2.18 Diagram of the different types of azimuth perspective projections, (**a**) centrographic gnomonic, (**b**) stereographic, (**c**) scenographic, (**d**) orthographic

In relation to the tangent point of the plane, the projections can be distinguished as

- polar, with tangent point in correspondence with the poles;
- meridian, with tangent point lying on the equator;
- oblique, with any tangent point.

2.9.2 Development Projection

In development projections, the points of the Earth's surface are transferred onto an auxiliary surface which is then laid onto a plane to produce the map. The most commonly used projection has a cylinder as auxiliary surface (Fig. 2.19).

The Mercator representation is a direct cylindrical projection analytically modified in order to make it conform.

Gauss conformal representation, also known as Mercator transverse projection (or Universal Transverse Mercator – UTM), although with a grid shape similar to the cylindrical inverse, is actually analytical.

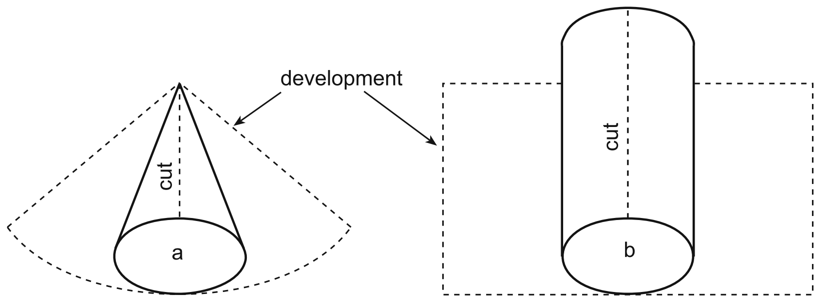


Fig. 2.19 Diagram of the (**a**) conic and (**b**) cylindrical projections

2.10 Examples of Cartographic Projections

2.10.1 Mercator Map

A Mercator map is derived from a central direct cylindrical projection, modified in order to make it conformal (Fig. 2.20). Using the ellipsoid as reference surface, the space between the parallels does not increase so evidently as in the central cylindrical projection.

A characteristic of Mercator map is to have meridians and parallels intersect at 90° angle, and a scale factor varies greatly depending on latitude.

Moreover, the use of this projection for areas far from the equator is not generally recommended, as the linear deforming module increases with latitude.

On a Mercator map, a straight line intersects all the meridians at a constant angle, being a line with a constant direction on the Earth: this way the lines with constant course, or *loxodromics*, look like straight lines, easy to trace, with an obvious advantage, for example, in navigational use.

2.10.2 Gauss Map

Gauss is the most commonly used projection in the world; it is similar to a cylindrical inverse, but derives from an orthomorphic cylindrical equatorial projection, originally suggested by Lambert, and later generalized by Gauss (1777–1855) who referred it to the ellipsoid (Fig. 2.21). During the last century, several scholars

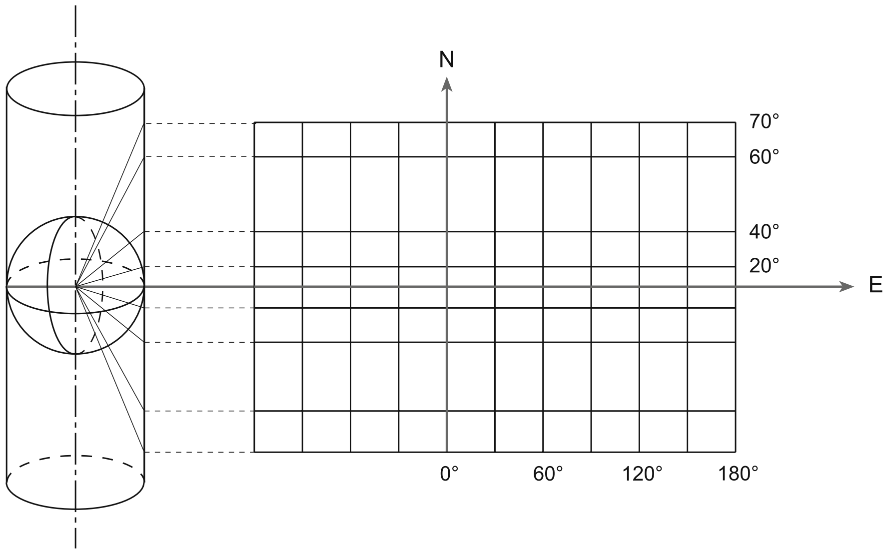


Fig. 2.20 Central direct cylindrical projection

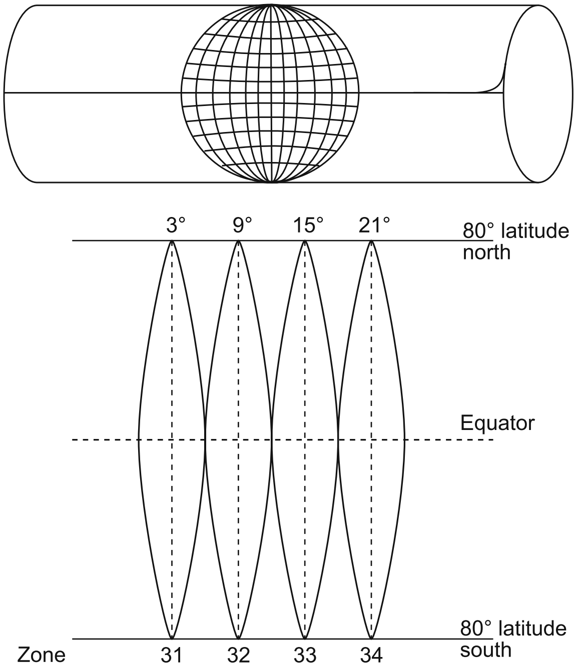


Fig. 2.21 Inverse cylindrical projection and definition of the zones from 31 to 34

introduced some modifications and reports on it, such as Boaga for Italy and Krueger for Europe. The UTM system is based on this projection.

The Gauss representation is conformal. The transformed central meridian and equator are straight lines that become the axes of the planar reference system and are called East and North. The transformed meridians and parallels are families of perpendicular curves, symmetric with respect to East and North axes.

In order to limit the linear distortions, it is necessary to represent an area which is subdivided into longitude bands, called zones. In the UTM system, the Earth has been divided into 60 zones, each 6° wide, numbered from 1 to 60 proceeding from West to East and starting from the Greenwich anti-meridian. The area that can be represented in a zone is a small portion of the Earth's surface and thus it is necessary to join more zones in case wider areas need to be represented. This representation system is called polycylindrical. Between two adjacent zones, there is always an area of overlap, in which the coordinates of both zones can be expressed.

Every zone is divided into 20 horizontal belts marked by letters and 8° wide (up to ϕ: $\pm$ 80°).

Every zone has its own planar reference system, which is independent as the central meridian changes from one zone to another.

Every zone has a false origin located 500 km East of the central meridian, in order to avoid negative coordinates, as at the equator the zone is about 666 km large, and the North coordinate originates at the equator (Fig. 2.22).

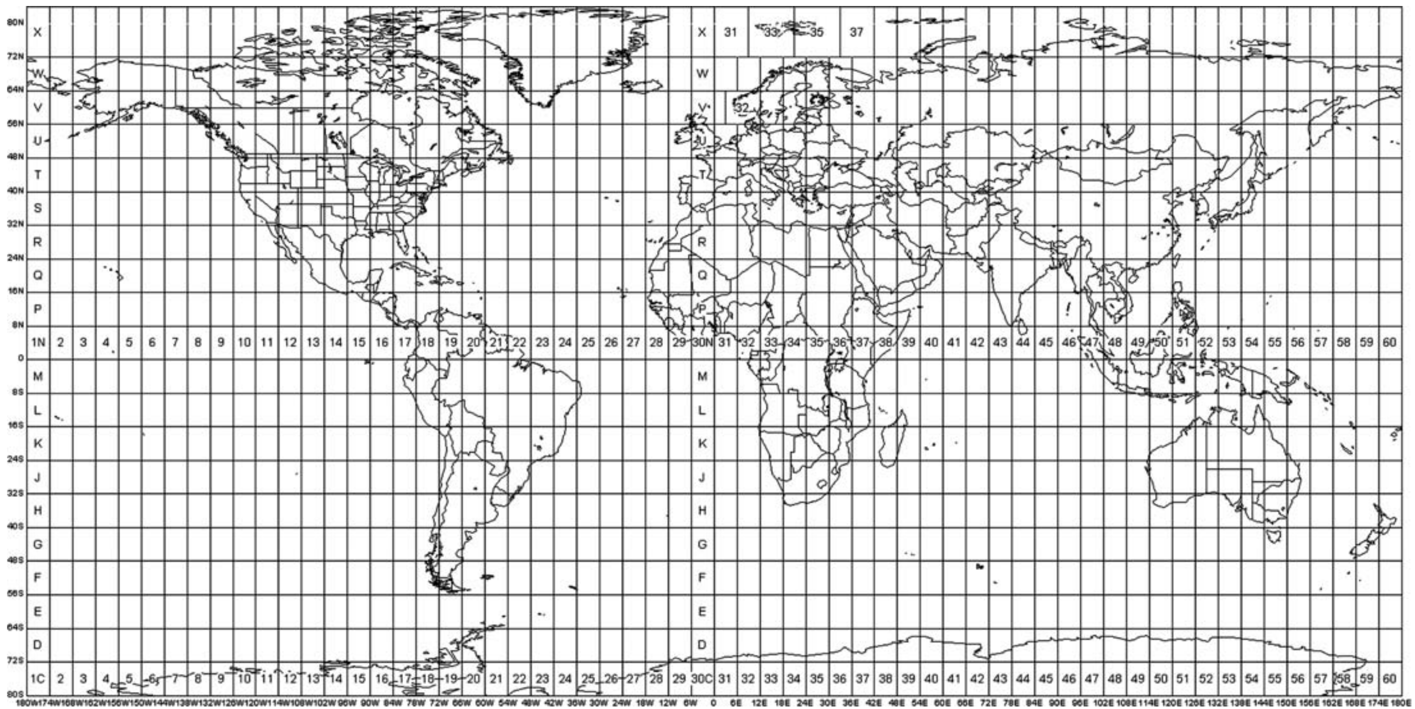


Fig. 2.22 Representation of the 60 zones each extending through 6° of longitude and of the 20 belts each extending through 8° of latitude (between 84° N and 80° S) of the UTM system. The zones are numbered consecutively from 1 to 60 beginning with the zone between 180° W and 174° W and continuing eastwards

Every zone is divided into squares 100 km as a side, marked by pairs of letters. Any point of the globe is bi-univocally determined by

- *two numbers*: zone;
- *three letters*: one for the belt, two for the square;
- *eight consecutive numbers*: four for the E and four for the N.

2.10.3 Polar Stereographic Projection

This is a pure perspective projection which is conformal. It is used to represent the polar caps, integrating the UTM cartography. In this representation, parallels are traced as circumferences, while meridians are transformed into straight lines. Parallels and meridians intersect at a 90° angle.

2.10.4 Lambert Conical Conformal Projection

The lambert projection is an analytical representation which can be derived from a direct conical projection, where the cone axis corresponds to the terrestrial rotation axis. The cone is tangent to a parallel, called the standard parallel, which is usually taken as the medium parallel of the area to be represented.

Developing the cone onto a plane, the map surface is defined within a circular sector. Meridians are represented on the plane by the cone generating line and thus by straight lines coming out from the homologous point of the cone vertex, while parallels are represented by arcs of circumference which have their centre in the meridians' convergence point (Fig. 2.23).

2.10.5 Earth Globe Projection: The Planisphere

The possible equivalent representations of the whole Earth's surface, also called planispheres, are numerous.

The globe or planisphere is an evocative and often artistic way to represent the Earth at a very small scale.

2.11 Reference Scale

When using a planisphere, the reference scale is the ratio between the measures taken among the points of its surface and the measures taken among the same points on the Earth. If the diameter of the planisphere is 1 m, the scale is 1:12,756,370 since this number is the diameter of the Earth in metres, supposing it has a spherical shape.

The scale of a map is the ratio of a single unit of distance on the map to the equivalent distance on the ground. The scale can be expressed in four ways: as a

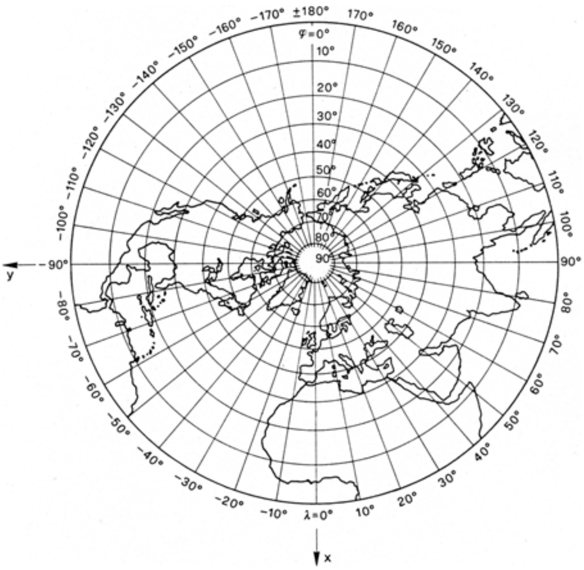


Fig. 2.23 Lambert central conic projection with stereographic polar projection of the boreal hemisphere

ratio, as a fraction, in words and as a graphical scale. Other secondary ways, like the *area scale* and the *relative scale,* are only seldom utilized.

2.11.1 Scale Factor or Scale of Reduction

If the scale of a map is formulated as the ratio between two distances, on the map and on the Earth, this relation is defined as scale factor or representative fraction. The scale factor is dimensionless, thus both terms of the relation must have the same units. The first term represents the distance on the map; the second represents the distance on the Earth. We define reduction scale, or scale, as the ratio between a length measured on the map and the corresponding one measured on the ground:

$$\mathrm{C} : \mathrm{T} = 1 : \mathrm{N}$$

where: C: length on the map; T: real length on the ground; N: scale denominator.

For example, in the scale factor 1:25,000, which can be written as a fraction: 1/25,000, 1 cm on the map corresponds to 25,000 cm on the terrestrial surface.

2.11.2 Graphical Scale

A graphical scale is a ruler with ground distances added, included in the margin of most maps. The graphical scale is used to measure distances on the map corresponding to the real distance between two points on the Earth. The distance on the map is marked on the edge of a sheet of paper, which is then placed over the graphical bar scale and the distance read.

This type of scale facilitates the reading of maps reproduced in different dimensions as it follows the enlargement or the reduction.

With the scale factor, this is not possible: the scale has to be recalculated any time there is a reproduction which modifies the dimension of the map.

2.11.3 Area Scale

The area scale, or unit of area, namely the square of the linear scale, is suitable to express the relation existing between an area on the map and the corresponding terrestrial surface. It is used for quick plotting, detail drawing, estimating the number of hectares in given irregular area and accurate measurement of land. It is often placed beside the graphic scale in the thematic maps to allow a quick computation of the surface.

2.11.4 Relative Scale

A relative scale refers to the size of the representation on an image as compared to the size of the object on the ground. In this case, the scale refers to the pixel size.

In orthorectified digital images and numerical cartography, within which the concept of scale is not conceived as in traditional cartography on paper, the scale is usually adopted as a *nominal scale*: this scale coincides with the scale of a traditional map with the same contents, level of detail and precision.

The term *relative* can also be related to the definition of scale. In describing the same spatial data, a geographer could refer to it as small scale, whereas an ecologist could refer to it as large scale.

Therefore, the terms small, medium, large scale, or smaller or larger, are often used in common language with reference to the comparison between different scale maps. This quite general language has often created confusion and misunderstanding.

In geomatics, the term *small scale* refers to a map which covers a relatively large surface of the Earth with the possibility of detecting few details.

The term *large scale* refers to a map with a high level of detail which covers a relatively small surface of the terrestrial surface.

Although no official limits have been adopted internationally, a commonly agreed classification of topographic maps is shown in Table 2.2.

Table 2.2 Map scale

Map scale (scale factor)	Scale definition
1:2500 or larger	Very large
1:25,000 or larger	Large
Between 1:50,000 and 1:100,000	Medium
Between 1:250,000 and 1: 7,500,000	Small
Above (geographic atlas)	Very small

2.12 Cartography in the World

All over the world, many institutions, companies and organizations are engaged in activities associated with maps and mapping. In general, they work at different levels, starting from the national (or even sub-national) one and reaching the continental and even global level. International experiences in the field of cartography have always been made difficult by the highly variegated standards, datums and reference systems used for map production; for example, when comparing cartographic products coming from neighbour states. Therefore, international initiatives in establishing supra-national references and standards have become increasingly important over time, especially with the increasing use of global positioning systems through satellite measure.

2.12.1 Cartography Projection in the World

Traditionally, since maps production is a sensitive aspect of power maintenance, administration and government, and also due to the necessity to better depict the territory using local reference systems, cartography has been a national matter, developed and managed through standards and methodologies that were not shared with other nation states. During the last 50 years, a new impulse has been developed for enhancing international cooperation and the supra-national establishment of reference systems, but the vast majority of cartographic production available nowadays shows a large number of both datums and projections, as the example of Table 2.3 shows.

2.12.2 International Reference Systems

2.12.2.1 Geocentric System WGS84

As GPS technology expands into the field of topography, a geocentric and global reference system has become needed, in addition to local reference systems already in use: this is the World Geodetic System, whose latest revision is WGS84.

WGS84 is currently the reference system used by the Global Positioning System (GPS), among others. It is geocentric and globally consistent within ± 1 m. Current geodetic realizations of the geocentric reference system family International Terrestrial Reference System (ITRS) are geocentric, and internally consistent at the few-cm level, while still being metre-level consistent with WGS84 .

The WGS84 originally used the GRS80 reference ellipsoid but has undergone some minor refinements in later editions since its initial publication. Most of these refinements are important for high-precision orbital calculations for satellites but have little practical effect on typical topographical uses.

The latest major revision of WGS84 is also referred to as ‘Earth Gravity Model 1996’ (EGM96), first published in 1996, with revisions as recent as 2004. This model has the same reference ellipsoid as WGS84 but has a higher-fidelity geoid (roughly 100 km resolution versus 200 km for the original WGS84).

Table 2.3 List of national map projections and datums used in countries worldwide

Country	Coordinate reference system identifier <country>_<datum>/<coord. system-proj.>	Geodetic datum	Projection
Argentina	AG_SAD69/GK	SAD69	Transverse Mercator (Gauss–Kruger)
Australia	AU_AGD66/84/UTM	AGD66/84	Transverse Mercator (UTM)
	AU_GDA94/UTM	GDA94	Transverse Mercator (UTM)
Brasil	BR_SAD69/UTM	SAD69	Transverse Mercator (UTM)
Canada	CA_NAD83/LAMBERT	NAD83	Lambert conformal conic
China*	CH_NANKING1935/GK	NANKING1935	Transverse Mercator (Gauss–Kruger)
Egypt	EG_EGYPT1907/MERC	EGYPT1907	Transverse Mercator
France	FR_ED50/EUROLAMB	ED50	Lambert conformal conic
	FR_RGF93/LAMB93	RGF93	Lambert conformal conic
	FR_NTF/FR_LAMB	NTF	Lambert conformal conic
Germany	DE_ETRS89/UTM	ETRS89	Transverse Mercator (UTM)
	DE_PD83/GK_3	PD83	Transverse Mercator (Gauss–Kruger)
	DE_42/83/GK_3	42/83	Transverse Mercator (Gauss–Kruger)
	DE_DHDN/GK_3	DHDN	Transverse Mercator (Gauss–Kruger)
	DE_RD83/GK_3	RD83	Transverse Mercator (Gauss–Kruger)
Great Britain	GB_OSGB36/NATIONALGRID	OSGB36	Transverse Mercator
Hungary	HU_HD72/EOV	HD72	Oblique conformal cylindric
India	IN_EVEREST/POLY	EVEREST	Polyconic
	IN_WGS84/UTM	WGS84	Transverse Mercator (UTM)
Italy	IT_ED50/UTM	ED50	Transverse Mercator (UTM)
	IT_ROMA40/EAST_WEST	ROMA40	Transverse Mercator (Gauss–Boaga)
Japan	JP_TOKYO1918/GS	TOKYO1918	Transverse Mercator (Gauss–Schreiber)
	JP_JGD2000/GS	JGD2000	
Morocco	MA_MERCICH/LAMB	MERCHICH	Cubic Lambert conic
Netherlands	NL_RD/DUTCH_ST	RD	Oblique stereographic
Romania	RO_S42(89)/TM_6	S42(89)	Transverse Mercator
	RO_S42(89)/ST1970	S42(89)	Oblique stereographic

Table 2.3 (continued)

Country	Coordinate reference system identifier <country>_<datum>/<coord. system-proj.>	Geodetic datum	Projection
Russia	RU_PULKOVO42/GK	PULKOVO42	Transverse Mercator (Gauss–Kruger)
	RU_PULKOVO42/UTM	PULKOVO42	Transverse Mercator (UTM)
	RU_PZ-90/UTM	PZ-90	Transverse Mercator (UTM)
South Africa	ZA_HARTEBEESTEHOEK94/UTM	HARTEBEESTEHOEK94	Transverse Mercator (UTM)
	ZA_CAPE/UTM	CAPE	Transverse Mercator (UTM)
Spain	ES_ED50/UTM	ED50	Transverse Mercator (UTM)
Switzerland	CH_CH1903+/CH_PROJECTION+	1903+	Oblique conformal cylindric
	CH_CH1903/CH_PROJECTION	1903	Oblique conformal cylindric
United States	US_NAD83/UTM	NAD83	Transverse Mercator (UTM)
	US_NAD83/SPCS	NAD83	Varying with States
	US_NAD27(83)/ALBERTS	NAD27(83)	Alberts equal-area conic
	US_NAD27(83)/LAMBERT	NAD27(83)	Lambert conformal conic

*Most up-to-date information

2.12.2.2 Unified European Reference System (ED50)

After the Second World War, as extra-national alliances were being formed, the need to realize a European Cartographic System for the newborn European Community members countries came up: this would be placed alongside to the national systems that all the countries had adopted. This objective was achieved by a general compensation of the European geodetic networks according to the following criteria:

- reference ellipsoid: international ellipsoid (HAYFORD);
- emanation centre: European medium orientation;
- origin of longitude: Greenwich.

As regards the emanation centre, the deviation from the vertical has not been eliminated there, but a residual deviation has been left such as to minimize the deviations at the borders of the area to be represented. This is how the medium European orientation (*European Datum* 1950 - ED50) became defined.

The new orientation of the reference ellipsoid generates the variation of all the represented points coordinates. After the network general compensation operations, the countries' trigonometric point coordinates assumed different values than in the national system (e.g. for Italy, Roma40). It is not possible to define analytic relations between the coordinates of a same point in those two systems (ED50 and a national one). Therefore, correction maps and tables both in terms of geodetic coordinates (ϕ, λ) and plane ones (E, N), called isotransition tables and maps, have been compiled to allow one to pass from one system to another with satisfactory approximation. The adopted representation is the UTM, according to what said above, to which the kilometric grid refers. UTM ED50 sheets have geographic coordinates according to the meridians and parallels transformed, but obviously there is no correspondence between the national sheets and the sheets of UTM ED50.

For example, the UTM ED50 kilometric grid is overlapped on the Italian national Gauss–Boaga reference system, so in the new IGM UTM ED50 production the Gauss–Boaga grid is preserved at the border of every map.

2.12.2.3 North American Datum (NAD83)

The North American Datum is the official datum used for the primary geodetic network in North America. In the fields of cartography and land use, there are currently two North American Datums in use: the North American Datum of 1927 (NAD27) and the North American Datum of 1983 (NAD83). Both are geodetic reference systems, but each is based on different measurements.

NAD27 is a datum based on the Clarke Ellipsoid of 1866. This Ellipsoid was created by a manual survey of the entire continent. The geodetic centre of NAD27 is a base station at Meades Ranch in Kansas.

As both satellite and remote sensing technology improved and were made available for civilian applications, it became obvious that the NAD27 approximations were not sufficiently accurate.

The North American Datum of 1983 was created to meet requirements for better accuracy and precision. It is based on the GRS80 ellipsoid; an ellipsoid derived from satellite geodesy.

- ellipsoid: NAD83
- a (major semiaxis): 6,378,137 m
- e^2 (eccentricity): 0.006694300229

A point having a given latitude and longitude in NAD27 may be displaced on the order of many tens of metres from another point having the identical latitude and longitude in NAD83. So it is important to specify the datum along with the coordinates. The North American Datum of 1927 is defined by the latitude and longitude of an initial point (Meades Ranch in Kansas), the direction of a line between this point and a specified second point, and two dimensions that define the spheroid. The North American Datum of 1983 is based on a newly defined spheroid (GRS80); it is an Earth-centred datum having no initial point or initial direction. NAD83 is the datum currently used for North American (U.S. and Canada) official cartography, commonly displayed using Lambert conformal conic projection.

2.12.2.4 Geocentric Reference System for the Americas (SIRGAS)

The Geocentric Reference System for the Americas (Sistema de Referencia Geocéntrico para las Americas, SIRGAS), created in 1993 with the support of the International Association of Geodesy (IAG), is the regional densification of the global International Terrestrial Reference Frame (ITRF). Its definition is identical to the International Terrestrial Reference System (ITRS). The reference coordinates are associated to a specific (reference) epoch, and their variation with time is taken into account by discrete station velocities or by a continuous velocity model, which comprises tectonic plate movements and crustal deformations. Realizations or densifications of SIRGAS associated to different reference epochs give rise to the same reference system and, after reducing their coordinates to the same epoch, are compatible at the millimetre level.

The SIRGAS geodetic datum is defined by the origin, orientation and scale of the SIRGAS system, and the geographical coordinates are derived by applying the parameters of the GRS80 ellipsoid.

The extension of the SIRGAS frame is carried out by national densifications of the continental network, which serve as local reference frames.

The first realization of SIRGAS (SIRGAS95) corresponds to ITRF94, epoch 1995.4. It is given by a high-precision GPS network of 58 points distributed over South America. In 2000, this network was re-measured and extended to the Caribbean, Central and North American countries. To account for this extension, the meaning of the acronym changed from the original *Sistema de Referencia Geocéntrico para América del Sur* to the current *Sistema de Referencia Geocéntrico para las Américas*. The new realization (SIRGAS2000) includes 184 GPS stations and

corresponds to ITRF2000, epoch 2000.4. The coordinate accuracy of these two realizations is about $\pm 3 \ldots \pm 6$ mm.

2.13 Transformation Among Reference Systems

Territorial information from different sources is often characterized by different reference systems or different types of coordinates. For this reason, there are specific software application packages that perform with good precision any kind of transformation from national systems to WGS84 data.

The complete treatment of this topic would require much more detail than this book can provide, so the reader is referred to specific publications.

When facing the problem of coordinate conversion from one reference system to another, it is important to consider that the ellipsoids are locally oriented and the commonly adopted altimetric measures, i.e. the orthometric height H referred to the geoid, is separated from the ellipsoid itself.

2.14 Map Classification

Maps can be classified according to different criteria and in function of the discriminative elements taken into consideration.

2.14.1 Basic and Thematic Cartography

The term *General Map,* used by Anglo-Saxon authors, is generally acknowledged to refer to a map. Thematic maps are a different class than general maps and are aimed at more specific or circumstantial purposes.

The different phases of the realization of a map are based on a sequence of logical processes applied to the knowledge of tangible or abstract phenomena.

According to common currents of opinion, maps can be classified according to

- content
- scale
- chronological facts.

As far the content is concerned, different types of maps can be distinguished as follows:

- *general maps,* or *reference maps*, or *topographic maps*: the ground elements are represented providing all the possible information, as far as the scale permits, without considering special phenomena that regard human activities or particular physical phenomena (i.e. determining climate conditions or factors). The purpose is to represent the territory while preserving a metrical proportion of the elements' size when possible or, alternatively, using a symbol;

- *thematic maps*: referring to a simple topographic base; the purpose is to represent the territory by surfaces or themes with uniform characteristics; the base of the thematic maps thus comes from information obtainable through general maps.

2.14.1.1 Topographic Maps

Maps representing morphological elements of the ground, hydrography included, using conventional symbols, are the foundation of the topographic maps category, which is extended to include the position of urban areas, roads, railway lines, political and administrative boundaries and other human-made infrastructure.

Topographic cartography is metrical; historically it has been produced using topographic and airborne photogrammetric surveys. With satellite acquisitions using high geometric resolution sensors is now possible to produce metrical maps from satellite data, until now mainly used to produce thematic maps only.

2.14.1.2 Thematic Maps

With regard to thematic cartography, difficulties related to the definition and classification are confirmed by the late suggestion of a unifying term: the expression *Thematic Cartography* appears in 1953, at least three centuries after the publication of the first cartographic documents nowadays defined as thematic. Anyway, this name cannot completely replace the heterogeneous presence of names like special, applied, inventory, analytical or synthetic cartography.

The thematic map is an instrument of landscape data representation, both geometric and statistical or demographic, where the different aspects of the territory are represented through qualitative and quantitative symbols laid on a topographic or geographic map. The production of a thematic map includes processes of information extraction from different original sources, such as already existent maps, in order to obtain a new product. For each thematic map, an appropriate legend is compiled, thus enabling the user to read one or more phenomena concerning the territory.

Two elements emerge from the wide range of definitions of thematic maps:

- the representation of one or more thematic classes, which introduce a characterization, is a specificity aimed at separation with respect to the basic cartography;
- the permanence of a basic reference which is strictly contiguous with the previous one is needed.

The thematic map, involving a different, more articulated knowledge and analysis level than general cartography, highlights the distribution of single geographic objects by knowing the internal spatial variations of a built space.

The realization of thematic maps is addressed not only to researchers in geographic, physical or anthropical fields, but many social sciences also use them to represent and analyse data with reference to their spatial position. With respect to

information representation using tables, the thematic map has greater capacity in terms of synthesis, visual effect and hypothesis of correlation among phenomena.

While the topographic map takes care of the metrical characteristics, the thematic map identifies the surface typology with precision. Thematic maps can be maps of altimetry, slope, exposure, land use, land cover, vegetation, lithology, erosion, urban development, land use modification, bathymetry (Fig. 2.24), etc.

Within thematic cartography, classes can be distinguished according to many parameters: characteristics of phenomena represented, representation means and techniques, theme categories.

*Analytical map*s provide information about the extent and distribution of one or more similar phenomena related to the characteristics of geographic space. These maps can represent *punctual* phenomena (e.g. well distribution), *linear* ones (e.g. the communication network) or *areal* ones (e.g. land use/land cover).

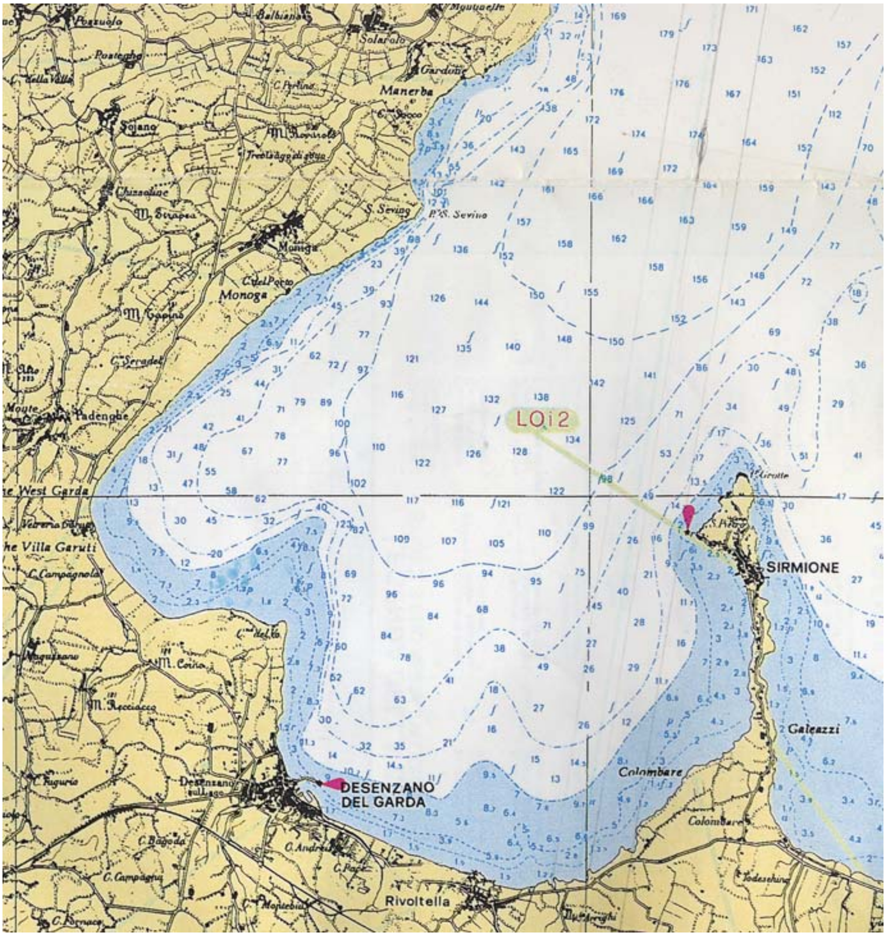


Fig. 2.24 Portion of the coastal map of the Garda Lake (Northern Italy) derived from the nautical map n. 862 Istituto Idrografico della Marina (IIM), Mercator projection

Synthetic maps provide correlations among more themes or define homogeneous areas according to specific unified analysis elements. These maps are being increasing by use in many applied fields related to land planning, operating with the overlay of different thematic planes in Geographical Information Systems.

Other expressions of thematic cartography come from depicting the instantaneous situation of a phenomenon, *static maps*, or from the cycle of its changes, *dynamic maps*.

According to the themes represented, maps can be distinguished in two categories:

- anthropic, representing themes directly referring to the human presence and activities, such as demographic density, population socio-economic characteristics, land use types;
- physical, to analyse natural phenomena, such as climate, geology, geomorphology, vegetation, fauna.

The symbology adopted for thematic cartography not only partially coincides with the general cartographic one but also has its own characteristics. As the thematic map is still closely related to its foundations, generally made up of physical (orography, hydrography, etc.) and/or anthropic (boundaries, road conditions, urban areas) elements, it is possible to exactly place the themes treated in the map and, at the same time, to note the correlations between the represented themes, distinctions and physical elements.

Moreover the thematic map can be

- *direct*, if finding the needed information is linked with the interpretation of primary sources that provide data collected by direct investigations, such as observations, prospecting, surveys, inquiries, photogrammetry and remote sensing;
- *derived*, if the information comes indirectly from secondary sources, i.e. by the use of already processed data and derivable from cartographic, statistical or bibliographical documents; for example the erosion map derived from the synthesis of information contained in the thematic maps of geomorphology, slope, exposure, etc.

As the orthophoto maps and orthoimages were introduced as the base on which the themes are overlayed, derived from aerial or satellite acquisitions giving an immediate view of the ground situation at the moment of acquisition and reporting, some references and numbers are needed to identify the classes.

2.14.2 Classification According to Scale

As the elements which can be directly represented are linked to their representation scale, obviously different scales require the adoption of different graphic standards (Table 2.4). In fact, reducing the scale factor increases the need to use conventional signs to represent ground elements of interest, which would not be representable

Table 2.4 Classification of thematic maps and reference scales

Thematic map	Scale
Cadastre	1:1000–1:4000
Project	1:2000–1:10,000
Land planning	Variable
Research	Variable
Atlas	> 1:1,000,000

elsewhere. If the maps have a kilometric grid and/or a geographic grid, with appropriate references, they are named *regular*.

Maps are classified, with reference to their scale, as

- *geographic*: from very small-scale planispheres 1:1,000,000;
- *corographic*: from 1: 1,000,000 to 1: 100,000 excluded;
- *topographic*: from 1:100,000 included to 1:5000;
- *technical*: higher than 1:10,000 scale.

2.14.3 Maps from Satellite

An image's geometric resolution limit is given by the image matrix cells size; enlargements over 2 pixels/mm generate the perception of a single cell as a discrete element (Box 2.2). For some applications where the spectral information is more important, it is possible to extend this resolution, producing an out-of-focus, pixellated image. When the images are reproduced and visualized by operating enlargements or reductions, it is more common to relate them to the ground by a graphic scale, which is better than the scale factor.

In any case, for convenience and comparison with traditional cartography the traditional scale factor is widely used. For studies dealing with the use of digital images, the representation of territorial information scales can be divided into three groups: global, regional or local.

A typical product that can be obtained through remotely sensed satellite images, undergoing a geometric correction and integrated with the insertion of some topographic elements and toponyms, is the *space-map*. This map is realized in very short time and is thus suitable for use in cases when updated information is more important than a detailed metric precision, i.e. for civil protection interventions.

2.14.3.1 Global Scale

This term refers to scales smaller than 1:250,000, used in studies at global level. In relation to the scale of intervention and to resolution characteristics, NOAA-AVHRR satellites are the most used. SPOT-Vegetation, Envisat-MERIS, Resurs-01, IRS-P3-WiFS, etc. are interesting for their potentialities. Possible applications are the study of the landscape's historical evolution, contribution to forecasting agricultural crops, monitoring phytosanitary condition of large cultivated and

wooded areas, production of small-scale thematic cartography, Vegetation Indices (VI), landscape units and land use analyses, thermal maps, agro-meteorological studies, etc.

Box 2.2 Relationship between the most frequently used satellite imagery, their applications and the most appropriate scales of cartographic restitution

<table>
<tr><th>Optical Scale 1:</th><th>2.5 Mil</th><th>1 Mil</th><th>500,000</th><th>250,000</th><th>100,000</th><th>50,000</th><th>25,000</th><th>10,000</th><th>5000</th></tr>
<tr><td colspan="10">Satellite/sensor</td></tr>
<tr><td>AVHRR SPOT Veg</td><td>GC LC G</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Resurs - 01</td><td></td><td colspan="2">GC LC G</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Landsat MSS</td><td></td><td></td><td colspan="2">GC LC G</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Landsat TM</td><td></td><td></td><td colspan="3">GC E LC G</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Landsat ETM+</td><td></td><td></td><td colspan="5">GC E LC G</td><td></td><td></td></tr>
<tr><td>SPOT XS, XI/IRS, LISS</td><td></td><td></td><td></td><td></td><td colspan="2">GC E LC G</td><td></td><td></td><td></td></tr>
<tr><td>SPOT Pan</td><td></td><td></td><td></td><td></td><td></td><td colspan="2">T LC G</td><td></td><td></td></tr>
<tr><td>IRS Pan</td><td></td><td></td><td></td><td></td><td></td><td colspan="3">T LC G</td><td></td></tr>
<tr><td>Ikonos, QuickBird, OrbView, DK 1-2</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td colspan="2">T LC G</td></tr>
</table>

<table>
<tr><th>Radar Scale 1:</th><th>2.5 Mil</th><th>1 Mil</th><th>500,000</th><th>250,000</th><th>100,000</th><th>50,000</th><th>25,000</th><th>10,000</th></tr>
<tr><td>Radarsat ERS 1-2, Envisat, Adeos, Aqua</td><td></td><td></td><td colspan="4">GC LC G</td><td></td><td></td></tr>
<tr><td>COSMO/SkyMed</td><td></td><td></td><td></td><td></td><td></td><td></td><td colspan="2">T LC G</td></tr>
</table>

GC: General Cartography (geographic and corographic maps)
LC: Land cover
E: Environmental thematic cartography
G: Geological cartography
T: Topographic cartography

2.14.3.2 Regional Scale

This includes scales between 1:250,000 and 1:100,000, useful for regional or basin level studies. The second- and third-generation satellite sensors for remote sensing that belong to this group are Landsat TM, SPOT-HRV/XS, IRS-1C/LISS, Landsat ETM+, MODIS, etc. Updated and synoptic land use/land cover thematic cartography obtained by image classification is a valuable tool for medium-scale analysis.

CORINE and AFRICOVER projects have been realized at this scale.

2.14.3.3 Local Scale

The use of satellite remote sensing always refers to smaller scales than those generally used for technical cartography (1:5000–1:10,000), which is related to the representation pixel size which, in normal contrast conditions, does not affect the continuous vision of the image when its side is smaller than 0.5 mm. With this limit, the 10 m HRV/Panchromatic SPOT allows a theoretical representation scale of 1:20,000. Possible applications are land use/land cover classification, topographic and thematic cartography updating, etc.

Airborne multispectral and most of all hyperspectral acquisitions allow thematic representation at detailed scales smaller than 1:5000, with the possibility of going beyond the graphic resolution limit due to the high spectral information content.

New perspectives arise with the availability of space images acquired by high-definition sensors which, having geometric resolution lower than 1 m and able to perform stereoscopic acquisition, bring the use of remote sensing in large-scale maps production closes.

Maps have been classified in time in different ways, some of which are reported in Box 2.3.

Box 2.3 Map classification systems used in the past

Classification system	Projection
Classification based on the geometric properties	Prospective for development epicylindrical or epiconical
Classification based on the geographic grid	Cylindrical mericylindrical conical central meri-conical polyphonic spherical merispherical star
Classification based on the deformation	Orthogonal or isogonics or conformal or orthomorphic equivalent or authalic aphylactic

2.15 Technology and Cartography: Numerical and Digital Cartography

Nowadays, cartography has to deal with the vast spread of information technology, which is becoming manifest in everything, not only in scientific fields. By retaining its philosophical characteristics of consistency, truth and readability, cartographic subjects are passing through an evolutionary phase intended to assure their wider use. This changes its form and, through the form, the way it is used. Cartography skills nowadays presuppose a control of the data, less and less hardcopy, becoming more numerical, while applications within Geographical Information Systems (GIS) become the privileged users of basic spatial data. Cartography is available to users in three forms: *traditional*, *automatic* and *numerical/digital*, a sequence that also describes its evolution.

2.15.1 Traditional Cartography

Traditional cartography is a ground representation framed into a coordinate system, organized in drawn tables, completed by a frame and appropriate parameters.

The information represented is of two kinds:

- *planimetric*: natural and artificial details of the ground are reproduced by projecting them onto the drawing plane;
- *altimetric*: contour lines and elevation of point define the ground altimetry. The information is conceptually separated from the planimetric data and it can be represented through appropriate descriptive symbols.
- The characteristics that turn a drawing into a map are also:
- *scale* (1:N), where N is the scale factor indicating the number of times the distance between two points is reduced on the map;
- *map key*, which gives the user the map interpretation key through the semantic evidence of the adopted symbols that refer to different types of lines, patterns, conventional signs, etc.

As stated above, a map is the product of the collaboration of disciplines like geodesy, topography, photogrammetry. It has to respect the three basic criteria of

- *consistency*, contradiction between reported types of information has to be excluded;
- *reliability*, the reported information must correspond to the reality;
- *readability*, the interpretation has to be unique.

These criteria are the foundation of any kind of cartographic production and are independent of the technical evolution and the details specifically required in each instance.

They are defined by compiling appropriate lists reporting technical rules to follow during the map's production in order to provide the customer with an a priori guarantee of the final product's quality.

2.15.2 Automatic Cartography

Automatic cartography is the link between traditional cartography and the numerical version. It arises from production needs and is included in the traditional cartography production process in terms of numerical data management aimed at the reproduction of geodata on a map; it represents the traditional cartography content in numerical, vector or raster form and allows its easy reproduction.

It does not have a numerically codified structure of the geometric elements represented. It turns the hardcopy information into digital data during photogrammetric projection or later by vector or raster digitization of existing hardcopy cartography.

2.15.3 Numerical Cartography

Numerical cartography preserves all the qualitative and metric characteristics of traditional cartography, giving a representation defined as follows:

- *data storage and structuring* are performed in numerical form on magnetic media of the represented objects' coordinates and their codification with respect to the typology;
- *data visualization* on video allowing management and hardcopy reproduction (plotter);
- unlike traditional cartography, the visualization can be realized at any scale. It is then fundamental to be able to distinguish between:
  - *representation scale*, which scale, can be any depending only on the software management adopted;
  - *nominal scale, uniquely* defined by the metric accuracies respected during the production phase and established in the list of technical rules.

The nominal scale is the largest scale at which a numerical cartography map can be reproduced while preserving its metric accuracy. The visualizations at scales larger than the nominal one allow better reading of the data, but do not determine an increase of data accuracy.

Numerical cartography is a mirror image of traditional cartography in philosophical terms. While the latter is a drawing that contains, in implicit form, coordinates defining an object, the numerical form is made up of an organized archive of coordinates containing their visualization in implicit form.

The way numerical data are organized allows access to the electronic archives, which can be done through programming logics typical of computers. Hence information can be objectively extracted and is not only based on deductive skills of the human user in charge of the analysis.

Since it is intrinsically numerical, numerical cartography is univocal in metrical terms as it eliminates all the subjective elements in measurement operations done on hardcopy drawings and the problems related to this deformation and degradation. The object codification system also allows a qualitative univocity.

Numerical cartography automatically performs classification, selection, statistics based on the objects' codification, geometry and topographic position. This makes the datum preferable for the realization and use in *Geographical Information System* (GIS).

With regard to information content, numerical cartography can be divided into

- *planimetric*: only the objects' planimetric coordinates (E, N) are stored;
- *plano-altimetric*: the planimetric information about the objects' positioning (E, N coordinates pairs) and the altimetric information about the ground, which preserves the form of contour lines and points of equal elevation (height), are stored separately;
- *3D*: all the objects are described by their three spatial coordinates (E, N, orthometric height). In particular buildings volumetric units are described by their altitude at the eaves and at the base, assuring a useful three-dimensionality in different applications, not least the propagation of electromagnetic signals, which is interesting for telecommunications field.

2.15.3.1 Numerical Data Format

Numerical representation establishes the definition of a reference conventional nomenclature:

- *object*, any natural or artificial element not separable any further (lake, building, etc.);
- *geometrical element*, or *geometric primitive*, the minimum representation element (point, broken line, polygon) with which a codification is associated;
- *entity*: complex object made up of one or more geometric elements.

As in the case of a traditional representation, the graphic element which represents the biggest difference is constituted by the curve line; in numerical cartography, this does not actually exist and is replaced by a broken line whose proximity among vertices has to be so dense as to ensure a difference between the real trend, *continuous curve*, and the represented one, *broken line*, less than the graphical error to the map scale. This concept lies at the base of the different techniques used in the restitution phase for the numerical representation of curve lines.

2.15.3.2 Information Content of the Numerical Data

In terms of planimetric content, numerical cartography has some innovative aspects in comparison with traditional cartography, mainly in relation to the representation philosophy based on a coordinate archive. This kind of archiving excludes the

possibility of deduction processes typical of the direct observation of a drawing. In traditional cartography, a road can be intuitively recognized by exclusion, as it represents an area not belonging to the surrounding buildings, and not by its internal codification. In the numerical case, aiming the smart computerized access to the data, implicit information has to be made explicit by an appropriate codification. Hence, unlike the traditional case in which the objects represented have a background that can be classified deductively, in the numerical case there is no background to allow objects emerge from a visual inspection of the map. The whole area has to be filled in order to leave no space for deduction. Thus road sections and nodes have to be represented through codified polygons or points.

The territory is divided into homogeneous surfaces which are

- *isolated*: all the surfaces that are delimited by closed broken lines (polygons) defined as separation lines (also shallow heights, like pavements) between artificial structures from a public space walking level;
- *services*: areas intended to have homogeneous use occupied by permanent infrastructures, like railway stations, airports, water treatment plants, power plants;
- *green areas*: areas intended to have homogeneous use occupied by permanent land cover types like parks, camping areas, etc;
- *open spaces*: agricultural areas, bare soil areas, quarries, landfills, etc.;
- *water bodies*: lakes, rivers, harbours, basins, etc.;
- *road sections*: road sections defined longitudinally by the objects delimiting them and transversally by virtual lines traced at distances that depend on the use of the data;
- *nodes*: areas resulting from the intersection of many sections (squares, etc.).

Considering the altimetric aspect, when present, four categories are described:

- *elevation of points*;
- *contour lines*;
- *points of equal elevation describing the planimetry of an entity*;
- *height at the eaves of volumetric units*.

The last two categories only concern numerical cartography. In the first case, the altimetric description is given by a set of four coordinates, *code*, *East*, *North*, *Altitude*, of each vertex of a geometric element involved in defining the represented entity. They are usually at the base heights and have to be coherent with respect to the ground average altimetric trend indicated by the points of equal elevation and the contour lines.

In case of heights at the eaves of volumetric units, the auxiliary altimetric information is expressed as a correspondence of points internal to the volume described, to which the altitude at the eaves is associated.

2.15.3.3 Code System

Objects, geometric elements and codification of entities are necessary to assure data tracing and query automatisms which allow us to easily use geographic data.

One commonly used system (Fig. 2.25) is the tree structure, defined by strings of eight numerical characters subdivided into four pairs to provide description levels, for each of the 99 (01–99) different possible categories.

Character pairs AA represent categories from 01 to 99 for each level so each entity can be identified univocally.

Some examples of categories are Railway communications (01), Road conditions (02), Buildings and infrastructures (03), Hydrography (04), Vegetation (05), Orography (06), Administrative boundaries (07), directly surveyed points (08).

Cartography updating and multi-level query result facilitated this data organization.

2.15.3.4 Data Organization

In terms of data storage, the basic idea is to realize archive structures so as to allow quick access and efficient query management. This need is typical of the numerical cartography with respect to the traditional one and is at the basis of GIS.

The management of archives of codified coordinates that represent objects and entities is realized by relational structures, typical of electronic archives, which link information contents by pointers and field identifiers functional to this operation.

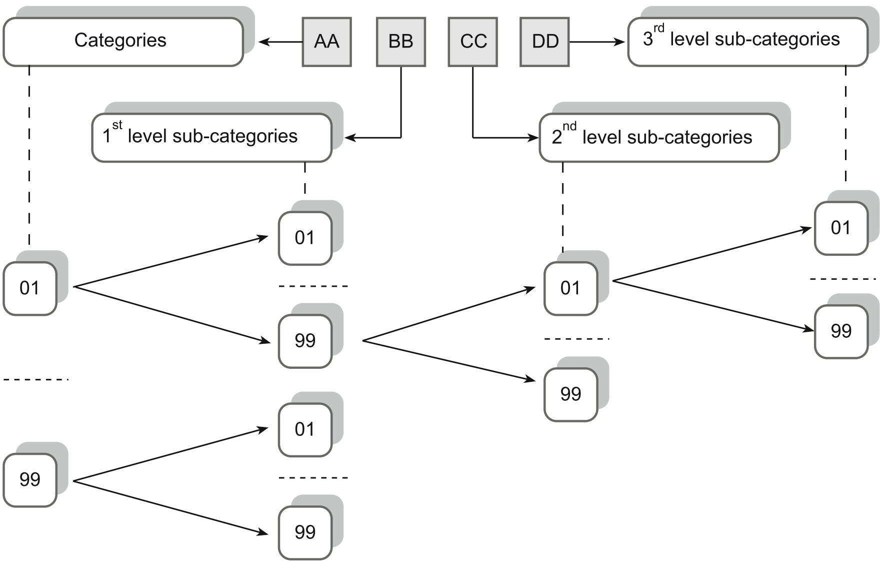


Fig. 2.25 Tree structure of the code of the entities for Numerical Cartography

Archive files are structured by record, object describer or entity, to which they refer. As the entities are represented by the coordinates (code, East, North and Altitude) of the vertices of the broken line delimiting them, each entity defines a record with a different length from the others. This problem affects access efficiency and thus query speed. The basic idea is hence to split the information through the realization of nested files containing constant length records, related among them by appropriate pointers. In particular the information will be split by defining a file that describes the entities and a coordinate's file linked to one another by pointers.

2.15.3.5 General Quality Criteria and Map Production Problems

This kind of numerical cartography requires, in the production phase, some controls for the solution of geometric incongruities which in the traditional case were automatically solved by the operator's direct intervention. So, for instance, the non-intersection among lines or the non-closing of a broken line that defines a polygon in the traditional case would be meaningless. In the numerical case, instead, if a broken line delimiting an area cannot close, this area is not recognized when an automatic query is launched.

Some editing procedures have been developed, which can solve the following types of situations:

- fusion of points representing the same vertex and many times assigned slightly different coordinates;
- non-intersection among lines;
- non-alignments;
- non-parallelism;
- non-squaring of right angles;
- non-continuity of lines conceptually continuous;
- non-congruence of altimetric data associated to contiguous entities.

Moreover, there are further geometric conditions and operational specifications to take into account when dealing with numerical cartography.

2.15.3.6 Production Methods

Three typologies of numerical cartographic data production are identified:

- direct field survey with topographic instruments;
- direct photogrammetric projection;
- digitization of existing cartography.

Though this division suggests a corresponding data quality classification, based on the various components that concur to define a numerical cartography quality, control networks precision, celerimetric survey quality, flight altitude, aerial

triangulation, editing quality, points collimation, etc., it is more appropriate to consider the categories listed above as a classification based on data generation:

- 1st generation, for direct survey;
- 2nd generation, for photogrammetric survey;
- 3rd generation, for the digitization of the previous two;
- 4th generation, for the direct numerical production during restitution phase.

Direct Survey

The direct survey not only is carried out by a tool set made up by an electronic or integrated theodolite (total station) for measuring the azimuth and zenith angles and the distances but could also be performed using an appropriate GPS instrumentation. It schedules the planning and realization of a general network and of the survey activity. Data are recorded in numerical format by the instrument and then transferred to a computer as fixed length files organized in codified records. Anyway, the codification and the file format have not yet been standardized.

Photogrammetric Survey

The direct photogrammetric survey is commonly realized in the same way as the classic photogrammetric projection. In this case (analogical, analytical or digital), plotters must have appropriate mass memories in which to store the large amount of coordinates and codes of the projected points and must complement automatic devices for measurements (encoders).

The difference with classic restitution is that the operator has to assign the appropriate code to all the projected entities. Dedicated visualization devices, or superimposition devices, allow, in the analytical and digital case, to overlap the already projected cartography in vector format onto video, so that the operator can have immediate feedback on his work.

Projection software, moreover, allows setting of the tolerance, the entity end and closure, the planimetric and altimetric closing, aimed to automatically solve some incongruence situations among those noted above.

There are four types of projected entities:

- punctual;
- linear or polylines;
- areal or polygon;
- text.

Plotters allow the automatic or semi-automatic generation of DEM (Digital Elevation Models) by collecting points that belong to regular grids with distribution and density characteristics defined by the operator. The result is an archive of X-, Y-, Z-coordinates corresponding to nodes of a regular grid.

Digitization of Existing Cartography

Third-generation numerical cartography production is certainly the most economical to realize. It includes a numerization process that transforms data on hardcopy maps into numerical data through digitizer devices.

This type of production, though, introduces some drawbacks:

- degradation of data precision with respect to the original characteristics of the hardcopy document:

– graphical error (0.2 mm) due to line heterogeneity and thickness;
– support deformation (up to 1–2% values and can be anisotropic);
– any parameters transfer error;

- data precision degradation with respect to the original one, transformed by numerization instruments:

– precision of the used digitizer;
– precision in the points collimation/restitution, related to the quality of the pointing device;

- considering its derived nature, it could refer to an obsolete hardcopy data source with insufficient updating;
- the produced numerical data is planimetric. It is always possible to integrate it later with elevation of points and contour lines, adding graphic layers.

Instruments for the Digitization

Digitization, or numerization, is the operation that turns the data on hardcopy into numerical data by means of digitizers:

- manual vector digitizers;
- automatic scanning raster digitizers;
- manual vector digitizers (Fig. 2.26).

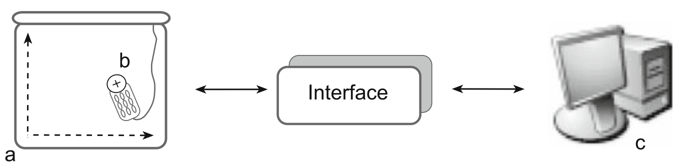


Fig. 2.26 Digitization system for vector numerization (**a**) digitizer, (**b**) cursor, (**c**) processor

They are constituted of three main elements:

- plain surface (table) where the map lies;
- pointing and acquisition device (cursor);
- interface connecting to a computer.

The table is constituted by a plastic support over which a grid of linear conductors, run through by electric power, is overlaid; the cursor indicates a solenoid whose centre is represented by the pointing cross that defines the position of the solenoid in respect to the absolute reference that is the table. An appropriate interface transfers the position information to the computer that records it. A keyboard designed to assign the codes completes the pointing device. These instruments' points positioning precision varies from 0.5 to 0.1 mm and mainly depends on the resolution (diameter and distance of the conductors in the grid).

Before performing this kind of acquisition, the operator has to orient the map, namely defining the planar roto-translation parameters with a four-parameter scale variation (most common transformation) to make the table coordinates (X, Y) coincide with the cartographic coordinates (E, N). This operation requires the identification of at least two control points on the map, their acquisition and the assignation of the corresponding cartographic coordinates. The number of points to be used in this phase (possibly belonging to the frame) is defined in the lists of rules as well as the highest residual differences acceptable after the orientation (generally < 0.3 graphical mm).

Raster Scanners

A raster digitizer, known as a scanner, is an automatic digitization system for which data, though numerical, are expressed as a matrix, like in a digital image (Fig. 2.27), and not as an archive of coded coordinates or vector mode.

In this representation, each element of the matrix generated during the scanning process represents the radiometric tone, or grey tone, correspondent to the graphic element that it represents. Thus black cells will refer to mapped entities, white cells to the background areas. The image matrix is generated by the following steps:

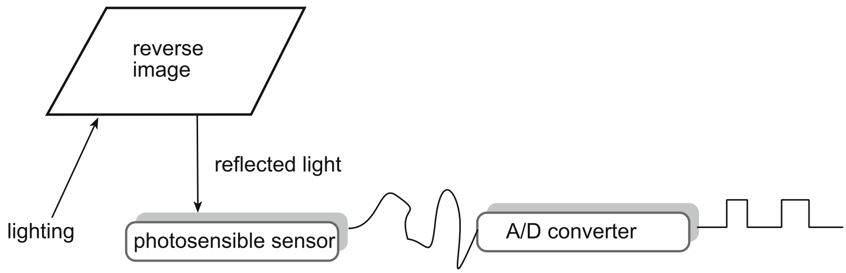


Fig. 2.27 Conceptual diagram of a scanner

illumination of the hardcopy support with a light source, acquisition (photo-sensitive sensor) of the luminous radiation reflected by the support itself and recorded signal analogical/digital conversion by a digital sampler. If the support is a colour one, the scanning is performed independently for the single representation bands RGB (red, green, blue) and consequently three independent matrices will be generated.

The scanning process is related to the resolution of the instrument, which defines the cell size in the image matrix, and according to the geometrical distortions generated.

The resolution is measured in dpi (dots per inch); the instruments on the market perform resolutions variable between 300 and 4000 dpi corresponding to cells' physical size ranging between 80 and 7 μm. The lower limit is fixed in dependence of electronic constraints on the signal-to-noise ratio.

The most common devices aimed at signal detection are the CCD (Charge-Coupled Device), electronic components (chips) made up by photo-sensitive elements organized in linear vectors or matrices and able to transfer the radiometric information through variations induced in the electric units. The signal recorded by a CCD is an analogical one. At the end of the digitization, a sampling device (analogical/digital converter) has to proceed to the transformation. This can be performed according to different modes:

- *binary digitization*: the recorded signal is represented by one bit only, this way determining the assignation of a value between 0 and 1 to each cell of the image matrix.
- *grey tone digitization*: the signal is represented by a higher number of bits (4, 8, 12) such as to ensure an image made up by a number of degrading grey tones depending on the number of bits employed (4 bits: 2^4 values, 8 bits:2^8 values, etc.):
- *colour digitization*: the generation of colour images can be obtained through 8-bit single scanning followed by assignation of a codified RGB palette, or by a scanning process characterized by separated RGB bands (24 bits) to be reconstructed in additive synthesis.

As the radiometric resolution, i.e. the representation bit number, increases, the amount of data to be stored increases too.

The process of image orientation also has to be carried out in the case of the scanning process, but this time after the acquisition, in order to align lines and columns respectively to the East–West and North–South directions. The transformation can still be a planar roto-translation with four-parameter scale variation, still with at least two necessary control points. Unlike the vector case, this operation is realized on a raster datum that involves all the cells, with reference to the entities mapped on the background, and requires a radiometric resampling operation that can alter the represented objects' geometry.

In order to improve the processing and to transform represented entities into vectors, further processing procedures can be carried out.

The instruments on the market, according to their use and performance, can be classified as follows:

- *DTP scanners* (DeskTop Publishing): addressed to non-cartographic applicative domains, they are the cheapest and most common product. They are planar and reach 1200 dpi geometric resolution. Usually they show such geometrical distortions and mechanical positioning instability to exclude an appropriate cartographic use.
- *photogrammetric scanners*: they were ideated within the photogrammetric domain and have very high geometric resolutions and mechanical positioning accuracy ranging between 2 and 5 μm. The present size limits their use in relation to the hardcopy datum to be digitized, as they can work with an A4 sheet size at maximum. They are planar instruments, too.
- *cartographic scanners*: they are generally rotating drum type, a fact which limits accuracy. They allow acquisition of size formats up to A0.

2.15.3.7 Data Transfer

The sharp conceptual and operative separation between the datum production and its use within a GIS requires facing the problem of transferring the numerical utilization, which is not yet univocally solved.

The production phase, as a consequence of production needs related to the instrumentation and to the operative modes, generates the datum in different *file working* formats that are often not managed by the information systems: these ones in fact, due to their internal structures and to the applicative domains to which they refer, tend to use different management file formats and surely different from those used in production.

This fact suggests defining a unified transfer file format so as to easily put the two productive and applicative/management segments in communication. For example, the main transfer files today adopted in Italy are the DIGEST (Digital Geographic Information Exchange Standard) and the NTF (National Transfer Format).

2.16 Map Reading

Geographic maps, and in particular large-scale topographic maps that are even more detailed, provide a faithful description of the characteristics of the represented territory. Their description is more detailed the bigger the elements drawn by conventional symbols.

In order to make clear the elements of the geographic maps' description, it is important to recognize the meaning of everything the map reports and to be able to interpret what the map represents from a visual point of view.

According to its definition, a geographic map is a plane, approximated, reduced and symbolic representation of a more or less extended portion of a terrestrial

surface. The symbols used, arbitrarily chosen by the cartographers in the past, have now been established by proper international conventional rules.

In modern maps, the different entities are represented as seen from above, according to a zenith view, not in partial or total perspective, as used to be the case in ancient maps. Maps, and in particular the topographic ones, show varied and considerable information; they represent existing physical or anthropic–geographic elements and imaginary elements too – like administrative, regional boundaries – different geographical grids, as well as information and data about the systematic, the editorial characteristics, etc., similarly to how information is reported on a printed book.

The elements reported on maps can be grouped as follows:

- systematic and editorial ones;
- geodetic–topographic, metrics and similar ones;
- natural landscape;
- anthropic landscape.

Cartographic symbols mainly reflect the last two elements. Its purpose is to describe the real world on maps, using conventional signs that are comprehensible to anyone and as quickly as possible. Some *conventional signs* have a direct relation with the represented object aspect (imitative symbols), both for their shape or for the organization of the elements that constitute them, while other ones, on the contrary, are independent of it.

Generally these symbols are always associated ones, as the landscape is mainly constituted by a complex of elements. The topographic representation of landscapes made by a single element, such as sea areas, big lakes, deserts, glaciers, can be considered exceptional. In flat alluvial zones, the represented elements are usually vegetation, hydrography, human activity elements, while in mountainous areas, together with the previous ones, there are others about orography, soil and rock aspects, exposures, etc. In the first case, there is a simple landscape, while in the second one the landscape is infinitely more complex as altimetric and planimetric aspects sum together; the map representing this latter situation requires more experience and good knowledge of representation techniques.

Legends were introduced in the XVI century, when two kinds of signs were used at the same time, together with writings and numbers, to represent, describe and denominate objects on a map:

- *symbols*: signs only in part independent of what they represent;
- *signs*: conventional signs completely independent of what they represent.

Usually the symbols are made by black graphical signs and, in polychrome maps, by colours with colour tones varying with symbol categories.

Simple graphical signs of landscape elements are points and lines which, properly arranged and combined, recall the real aspect of what is on the territory, making things easier for those who analyse the map.

For a long time, maps have been not only graphical products but real artistic compositions. With the modern representation techniques, graphical aspects have been partially preserved but artistic ones have been lost.

2.16.1 Elements of the Natural Landscape

Symbols used to represent elements and shapes of natural landscapes are much more numerous than geodetic and topographic ones and can be grouped into the following categories:

- orography;
- continental hydrography;
- marine hydrography;
- natural vegetation.

The last three groups of elements can be considered as one group including entities characterizing the planimetry of a region.

Orography is a non-flat region relief and can be represented using several techniques and systems variable in function of map type and scale. Lots of these systems are obsolete but have been widely used in the past, like

- altimetric zonation;
- dash, stroke or line;
- herring-bone system;
- hill shading;
- hatching;
- contour line system.

Since methods and instruments reconstructing and visualizing 3D models of the territory also in a dynamic way and with different angles of view (see Chapter 9) have been developed, the systems described above are valid for reading the already existing cartography and for the production of cartography that still relies on them.

The shapes that contribute to represent orography can be grouped in different ways, for example, in relation to the type of phenomena they represent (glaciers, carsism and volcanic phenomena, etc.) or to the specific objects (rocks, debris, peaks, etc.).

2.16.2 Elements of the Anthropic Landscape

Natural landscape elements and shapes are represented with particular symbols, which reflect human activities and most of its products. They have transformed the original aspect of the landscape and contributed to produce a new, anthropic one. The bigger the map scale, the bigger the number of elements in the anthropic

landscape that can be represented. Some are independent and others correlate to those described above. They can be grouped as follows:

- human settlements and various buildings;
- communication lines;
- irrigation systems;
- agricultural land use, natural land cover excluded;
- quarrels, mines and similar;
- political and administrative boundaries, etc.

2.16.3 Generic Nomenclature

Besides the elements relative to the soil's physical aspect and to human activity, all the maps types report various kinds of different importance scripts, related to what is represented; the map without these scripts is called mute.

Until 1840, the toponyms used to be transcribed without any particular rule, then it was agreed to use special graphic signs for different groups of toponyms. This way the scripts got a symbolic meaning, as it is possible to understand the toponyms' importance or their differential relations in relation to the way they are realized (regular, italic, capital letters, bold, font size, etc.).

Toponyms of topographic maps, especially large-scale ones, often have intrinsic meanings, including for example a morphologic meaning. Thus, the toponym analysis is often useful in relation to the numerous cases in which a simple common name has become a toponym to which more than one place is associated. A complete toponym contains itself a reference to a particular aspect of the surface.

2.17 Summary

The purpose of cartography is to represent real geographic objects on maps, providing both a punctual and general knowledge of the territory, establishing spatial relationships among the reproduced objects and producing tools used in planning and management of territory. The origins of this science go back to very ancient times, passing through the Egyptians and Greeks, the Renaissance in the Mediterranean Basin, to the present day. One of its main issues since antiquity has been to give a 2D representation of the 3D Earth's surface; a crucial step in this process is the definition of a reference surface (which is the subject of Geodesy) that better approximates the Earth's surface in each portion. Once the *geoid* and the *ellipsoid* were defined as two different reference surfaces, the second one, responding to geometrical laws, was used to define several reference systems, both local and global. In relation to a reference system, several coordinate systems can be defined through mathematical models (DATUMs). The aim of a coordinate system is to describe the position of any point in respect of the chosen reference system.

Besides locally valid reference systems, some reference systems are recognized worldwide, like the WGS84 (*World Geodetic System 1984*), or at European level, the ED50 (*European Datum 1950*). The correspondence between local and global reference systems is not direct, and the coordinate conversion from one system to another is required. For this purpose, specific software performing with acceptable precision transformations from national, ED50 and WGS84 data have been conceived.

The production of maps is realized by projecting the reference surface of interest onto the map plane; this step produces some deformations. Maps are classified according to the type of deformation introduced by analytical transformation, or according to the characteristics of the projection or of the auxiliary surface onto which it is projected. Further parameters used to classify maps are the content of the map (general maps/thematic maps) and the scale (geographic, corographic, topographic, technical). The scale factor is a basic element of the map which gives an idea of the entity of the reduction of the map and affects the amount of detail that can be directly represented on the map itself.

As informatics is spreading in many sectors of science, cartography too has been influenced by computer technology. If traditional cartography is realized on paper boards, digital cartography is based on numerical format data that are visualized on a screen and allows the data to be dynamically used in Geographical Information Systems.

To be used, both traditional and digital maps need to be interpreted. In fact a geographic map is defined as ‘a plane, approximated, reduced and symbolic representation of a more or less extended portion of terrestrial surface’; according to its scale, any map reproduces many elements of the represented surface, which are expressed as symbols codified by a legend. Both natural and man-made elements are symbolically depicted, and toponyms are reported.

Further Reading

Fenna D., 2007, Cartographic Science, A Compendium of Map Projections, with Derivations. CRC Press, Taylor & Francis Group, pp. 504, ISBN: 978-0-8493-8169-0.

Maling, D.H., 1992, Coordinate Systems and Map Projections, 2nd ed.. Pergamon Press, Oxford.

McDonnell, P.W. Jr., 1979, Introduction to Map Projections. Marcel Dekker Inc., New York, pp. 174; 2nd ed., 1991. Landmark Enterprises, Rancho Cordova, California, pp. 198.

Richardus, P., Adler, R.K., 1974, Map Projection: For Geodesists, Cartographers and Geographers. North-Holland Publishing Company, Amsterdam, The Netherlands.

Robinson A., Sale R., Morrison J., 1978, Elements of Cartography, 4th ed. John Wiley & Sons Inc., New York.

Robinson A.H., Morrison J.L., Muehrcke P.C., Kimerling A.J., Guptill S.C., 1995, Elements of Cartography. John Wiley & Sons, New York.

Bibliography

AA.VV. 1980, Grande Atlante Geografico De Agostini, Ed. De Agostini, Novara.

Avery T.E., Berlin G.L., 1985, Interpretation of Aerial Photographs, Chapter 6: Planimetric and Topographic Mapping. Burgess Publishing Company, Minneapolis, Minnesota, USA, pp. 115–140.

Bertin J., 1977, La graphique et le traitement graphique de l'information.
Bugayevskiy, L.M., Snyder J.P., 1995, Map Projections: A Reference Manual. Taylor & Francis London.
Canters F., Decleir H., 1989, The World in Perspective—A Directory of World Map Projections. John Wiley & Sons, Chichester, England.
Denègre J., 1988, Thematic Mapping from Satellite Imagery, An International Report, International Cartographic Association, Ed. Elsevier, London and New York.
Dorling D., Fairbairn D., 1997, Mapping: Ways of Representing the World. Prentice Hall, Pearson Education, Harlow, England.
Eckert M., 1907, Die Kartographie ale Wissenschaft, in Zeltschr, der Gesell, fur Erdkunde, Berlino, pp. 539–555; ID, On the Nature of Maps and Maps Logic, in Bulletin of American Geographic Society, New York, pp. 344–351.
International Cartographic Association, Commission II, 1973, Multilingual Dictionary of Technical Terms in Cartography. Franz Steiner Verlag, Germany.
Petchenik B.B., 1979, From Place to Space: The Psychological Achievement of Thematic Mapping. American Cartographer, Falls Church.
Rouleau B., 2000, Méthode de la Cartographie. CNRS Editions, France.
Snyder J.P., 1984, Map Projection Used by the U.S. Geological Survey, Geological Survey Bulletin 1532. US Government Printing Office, Washington, USA.
Snyder J.P., 1993, Flattening the Earth—Two Thousand Years of Map Projections. University of Chicago Press, Chicago.
Snyder J.P., Voxland P.M., 1994, An Album of Map Projections, U.S.G.S. Professional Paper 1453. Government Printing Office, Denver, USA.

Chapter 3
Elements of Photogrammetry

Photogrammetry is a technique that allows the measurement of an object without touching it. Measurement can be performed in two and three dimensions (2D and 3D) exploiting both photograms (analogical images) acquired by traditional photogrammetric cameras and digital imagery. Although photogrammetry was born for architectural survey, it can be considered the first remote sensing technology based on the acquisition of objects' geometric properties from photographic images. Nowadays it is widely used in topographic aerial survey and mapping, and for military purposes.

The term Remote Sensing (RS) is often used instead of photogrammetry. RS is a young term, which was originally confined to working with aerial photographs and satellite images. Today, it is in synergy with photogrammetry, although it is still associated rather with image interpretation.

Basic elements of traditional and digital photogrammetry are described below. The introduction gives an overview on experiences progressively matured since the beginning of the 1800s in photography, later in photogrammetry in the nineteenth century and, with the development of aeronautics, in aerial photogrammetry.

The 3D coordinates of points on an object are determined by measurements made in two or more photographic images taken from different points of view: an appropriate processing of such images determines the artificial stereoscopic vision. The identification of common points on different images of the same object defines a line, or ray, of sight geometrically linking the image point to the corresponding object-point. The intersection in space of two or more of these rays determines the 3D position of the point (according to the *forward intersection* principle).

3.1 Milestones in the History of Photography

Photography (*drawing with light*) was officially born in 1839 when the French Academy of Sciences announced that Louis Jacques Mande-Daguerre had succeeded in improving the previous studies developed by Joseph Nicèphore Nièpce (1765–1832). Daguerre succeeded in impressing for the first time objects of the real

M.A. Gomarasca, *Basics of Geomatics*, DOI 10.1007/978-1-4020-9014-1_3,

Fig. 3.1 First photography (18 cm × 24 cm) conserved in Austin, University of Texas (USA) since 1964. It is an image of a French landscape shot in 1826. Nicéphore Niepce obtained the image with Judea bitumen spread on a silver plate for heliography after an exposure time of 8 h. The photograph shows a view from his home window (Le Gras, Chalon sur Saone). This is the only preserved image achieved by Niépce with a camera obscura that is representative of this step of his research

world on a photo-sensitive film with an exposure of 30 min in comparison with the 80 h needed for the first tests of eliography (*drawing with Sun*) already effected by Nièpce in 1821, and the 8 h of the first true photography (Fig. 3.1).

For his first experiments, Nicéphore Niépce positioned at the back of a *camera obscura* a sheet of silver salt-coated paper, known to blacken with daylight. In May 1816, he produced the first image of nature: a view from a window. It was a negative and the image vanished because in broad daylight the coated paper becomes completely black. He called these images retinas.

The term *photography* is derived from the Greek *photos* (light) and *graphein* (to draw). This word was first used by the scientist Sir John F.W. Herschel in 1839. It is a method of recording images by the action of light, or related radiation, on a sensitive material.

The first element of a photo camera, known by the ancients, is the effect of illuminating an object located between a light source and a dark area with only a pin hole opening: an inverted (upside down) image of the illuminated area is produced on a flat surface beyond the dark area. As early as the 1400s, it was documented that replacing the hole with a lens would produce a crisper, clearer image. Leonardo da Vinci (1452–1519) described with precision the phenomenon, already known by the Arabs, in one of his manuscripts.

This technology, called *camera obscura,* was often used by artists to sketch objects more quickly and ease the difficulties of depth perception. The image was

allowed to be projected on a piece of paper inside a dark box and the artist would trace outlines of the projected image.

The second known element of a photo camera was the existence of materials capable of permanent change when exposed to light. These light-sensitive chemicals were experimented with for centuries but were not used to coat a flat surface until very recently.

A number of light-sensitive materials were tested to capture the image from the camera obscura, but the first successful permanent photograph is usually credited to Louis Daguerre. That picture, captured on a silver-coated sheet of copper, using his positive image Daguerreotype process, is entitled The Artist's Studio and is dated 1837. It was fragile and difficult to reproduce.

While the details of this process were made public, in 1839, other artists and scientists discovered additional photographic imaging techniques. William Henry Fox Talbot's Calotype process used light-sensitive paper and produced a 'negative image' that could be used to create positive prints.

At the beginning, the problem was that due to the long exposures of the photosensitive material, only static objects could be photographed. When lenses and thus optics placed over the films were developed, it was possible to manage the exposure time according to the aperture, i.e. the shutter, regulating the flux of light.

The photographers' aim has always been the faithful reproduction of the real object in the final image in terms of both resolution and chromatism. The first acquisitions were in black and white, and in 1861 film able to record colours of the photographed scene was produced.

The origin of colour photography goes back to the studies of the Scottish physician James C. Maxwell, in 1855, giving a philosophical explanation of additive colour synthesis, already figured out by Aristotele (384–322 B.C.) and discussed by Thomas Young in 1802 in the theory of human vision which assumed the different sensitivity of the retina to blue, green and red.

Maxwell also gave contributions to the topics of optics and colour vision, being credited with the discovery that colour photographs could be formed using red, green and blue (RGB) filters. He photographed a tartan ribbon three times, each time with a different colour filter over the lens. The three images were developed and then projected onto a screen with three different projectors, each equipped with the same colour filter used to take its image. When brought into focus, the three images formed a full colour image (Fig. 3.2).

If the vertices of an equilateral triangle correspond to the three colours RGB and admitting that colour saturation decreases moving from the vertex, as the colour combines with one or both of the other colours, each point in the diagram can be assigned RGB coordinates, in this way establishing exact combinations of red (R), green (G) and blue (B).

In the vertex of red, saturation is maximum for red, while green and blue are absent:

- R: 255,0,0

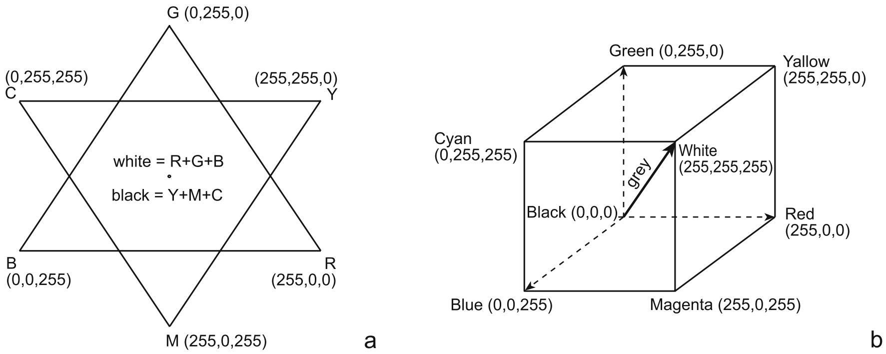


Fig. 3.2 Definition of the colorimetric space (*Red, Green, Blue*: RGB) used in the colour theory developed by Maxwell; (**a**) triangles of the additive model and (**b**) colour cube. The chromatic distribution depends on the level of saturation of the colours that depends on the luminosity. Examples with radiometric resolution of 8 bit: 256 levels

In the vertex of green, saturation is maximum for green, while red and blue are absent:

- G: 0,255,0

In the vertex of blue, saturation is maximum for blue, while red and green are absent:

- B: 0,0,255

White light coordinates correspond to the centre of the equilateral triangle, which is the intersection point of the heights, these corresponding to the medians and bisectors, equidistant from the vertices.

Thus Maxwell introduced the concepts of hue and saturation. Hue is represented along the sides of the triangle, where saturated colours corresponding to the mix of the two primary lights are located at the two adjacent vertices. At the medium point of each side, there are the respective *complementary* or *subtractive colours*, yellow, magenta and cyan, according to the following scheme:

Red + Green + Blue = White
Yellow + Magenta + Cyan = Black
Red + Green = Yellow
Red + Blue = Magenta
Green + Blue = Cyan

Going from the vertices of the triangle towards the white (centre), the triples of values progressively decrease in saturation.

Later A. Mansell ideated a colour classification system based on a set of coloured tags founded on five basic colours: red, yellow, green, blue, magenta and their

combinations, explaining that the colour is described by three units, or chromatic variables:

- *hue*: related to the distinction among different colours based on their wavelength;
- *brightness* (or intensity, or value): for each colour it is possible to distinguish different shades from the lighter to the darker. For example, red rose at midday in the sunlight and at twilight;
- *saturation* (or chroma): the relative purity of the colour, i.e. the presence of only one colour without the influence of any other.

The step from ground photography to aerial photography was made by G. Félix Tournachon, called Nadar, one of the first photographers who guessed the potentialities of aerial acquisition in 1856 equipping an aerostatic balloon to perform the first aerial stereoscopic acquisition of Paris. Nadar was a very popular artistic photographer of that period especially for his portraits, characterized by a strong realism, sometimes exaggerated, and also for his progressive spirit, open to technological innovations and those enterprises requiring some audacity and a bit of exhibitionism. He is said to have carried out some acquisition from aerostatic balloons the day before the battle of Solferino, in 1859, in order to spy on the Austrian army, the very first example of aerial military reconnaissance.

Even carrier pigeons were ingeniously used, as innately skilled to fly along the geodetic, i.e. the shortest way between two points, and to come back to the starting: a photographic camera for acquisitions for military purposes used to be fixed to their body (Fig. 3.3).

Photogrammetry, such it is known today, comes from photography; it revisits the photographic representation in metric terms (Box 3.1). The description of the object must include also its dimensions.

Fig. 3.3 Pigeon shooting system using a light photographic camera with scheduled timer every 30 s

Box 3.1 Reference year, scientist, and scientific discovery in the fields of photography and photogrammetry

Year	Scientist	Discovery
1630	G. Desargues	Formulation of the theory of the perspective
1727	J.H. Schulze	Discovery of the sensitivity of the silver nitrate exposed to the light obtaining for a very short time an image
1759	H. Lambert	Publication of the theory of the perspective *Frei Perspective* Introducing the principles of the inverse perspective and of the spatial intersection of the homologous perspective rays, fundamentals of the photogrammetry
1726	M.A. Kappeler	Survey of the Pilato mount applying the perspective theory
1807	W.H. Wollaston	Design of the Camera Lucida (Lucida is Latin for *light*), reflecting prism which enabled artists to draw outlines in correct perspective. Instrument overlapping a territorial image onto a geographic map
1826	N. Niepce	First photograph with 8 h of exposition Camera obscura, a device which reflected a scene onto a canvas that artists then traced
1839	L. Mandé Daguerre	Development of the first latent image formed on silver salts (iodide) He invented the process to permanently capture images on light-sensitive materials Others rapidly made improvements to the photographic process but Daguerreotypes remained the principal medium of photography until the 1850s. Daguerreotypes consisted of a silver-coated brass plate exposed to iodide vapour. The plate was then exposed to light and developed using mercury vapour. The image was fixed using sodium thiosulphate and was mounted under glass to protect it
1839	Arago	Presentation at the Academy of Sciences in Paris of the photography potentiality for metrical use
1840	F. Voigtlander	Collaborating with the mathematician J. Petzval designing the first lens with maximum aperture of f/3.7. To use the lens, he produced the first all-metal photo camera which was also the first to have rach-and-pinion focusing
1840	H. Fox Talbot	Negative–positive process reproducing photographs
1840	D. Brewster	Stereoscopic viewer
1841	A. Claudet	Application of the stereoscopic photographic techniques
1851	F. Scott Archer	Procedure of the humid colloids to preserve photographs for 40 years
1850	A. Laussedat	Fundamentals of the photogrammetry, defined as iconometry (icon [Greek] meaning image, -metry [Greek] which is the art, process or science of measuring)
1858	F. Tournachon detto Nadar	First aerial photograph from a fire-balloon
1859	A. Laussedat	First photo-theodolite: a theodolite equipped with a photo camera with four fiducial marks
1860	A. Maydenbauer	Large format photographic camera used in monuments survey, creating the first photographic architectural archive (Berlin 1883)

Year	Scientist	Discovery
1860	I. Porro	First photo-goniometer able to measure the photographic lens distortion,
1860		Use of aerostatic balloons for acquisition of aerial photographs during the American civil war
1871		Gelatine silver iodide glass plates, later improved (1890)
1876	W. Jordan	Official use of the term photogrammetry
1876	G. Hauck	Definition of principles of the ‘analytical photogrammetry’
1888	G. Eastman	Metric film used in a portable Kodak photo camera
1889	D. Koppe	First written document concerning photogrammetry
1895	Deville	First instrument for stereoscopic vision of photographs
1896	D. Koppe	New version of the photo-goniometer according to the so-called Porro–Koppe principle
1901	Karl Pulfrich	First photo-comparator, allowing a diffuse use of the analytic photogrammetry
1903	Fratelli Wright	Airplane
1908	E. Von Orel	Stereo-autograph, first analogical plotter for the direct restitution
1908	Theodor Scheimpflug	Considered as the initiator of aerial photogrammetry, succeeding to apply the photogrammetrical principles to aerial photographs (model orientation, strip orientation, restitution, rectification, etc.)
1923	Ermenegildo Santoni	Introduction of the stereo-planigraph, plotting instrument with optical–mechanical projection
1957	U. Helava	First analytical stereo-plotter, produced in cooperation with OMI and Bendix
1972	G. Inghilleri	Introduction of an innovative analytical stereo-plotter, jointly with S. Dequal, produced by Officine Galileo, Italy
1981	T. Sirjakoski	Introduction of the digital photogrammetry
1988	Kern	First digital stereo-plotter

3.2 Milestones in the History of Photogrammetry

The initiator of this discipline is traditionally considered the French colonel A. Laussedat: in 1859 he presented a report at the Committee of the Academy of the Sciences in Paris about the possibility to define points’ coordinates of an object based on the spatial intersection of projective rays referred to a pair of photographs.

At the same period, A. Meydenbauer carried out the first photogrammetric survey of buildings.

In 1901 Pulfrich gave the first input to the history of stereo-photogrammetry leading E. Von Orel to build the first stereo-autograph in 1909, an instrument able to continuously retrieve planimetric lines and contour lines from terrestrial photograms.

Later studies brought W. Bauersfeld in 1923 to apply the concept of the stereo-autograph to retrieve aerial photograms; he called the new instrument the *stereo-planigraph*.

Later technical development allowed the building of more efficient photogrammetric cameras and plotter instruments; thanks to these, supported by the theoretical research by characters like S. Finsterwalder and O. Von Gruber, photogrammetry finally became an accurate and cheap method for restitution.

The revolution brought about by the introduction of computers also involved photogrammetry, which, after changes in its operating methods, paradoxically went back to its original numerical character (the first photogrammetric surveys had numerical format and manual calculation). Computers progressively replaced the plotters' big optical–mechanical components, moving the approach from analogical to analytical. The next step from analytical to digital was possible thanks to the introduction of scanning image acquisition techniques.

At the same time, optimization numerical techniques introduced the aerial triangulation in block orientation of multiple photograms using a smaller number of footholds.

Nowadays the situation is transitional: classic photogrammetry is more firmly moving towards the digital, for generating images and the whole restitution process.

3.3 General Concepts

Photogrammetry is a technique devoted to the measurement of objects; its basics are related to the rigorous exploitation of the properties of the artificial stereoscopy easily obtainable by acquiring pairs (or more) of images of the same object from different points of view (Box 3.2).

Box 3.2 Definitions of products derived from aerophotogrammetry

Product	Definition
Rectified image or photoplane	Central projection of the original image with the same geometric characteristics of an orthophoto. This process can be adopted only in case of slant (oblique) acquisition of plane objects (elements)
Photo-mosaic	Mosaic of rectified images
Orthophoto or orthophoto-plane	A photo-map (where the objects are represented in their correct plane position) obtained by an analytical transformation of the acquisition perspective into an orthogonal projection
Orthophoto-mosaic	Mosaic of contiguous orthophotos
Digital orthophoto	Images with correct spatial positioning of the cells (pixel) by an analytical transformation of the acquisition perspective in orthometric perspective

The products of the photogrammetric process can be

- *numbers*: groups of 2D or 3D coordinates defining the geometry of measured objects points with respect to the adopted object reference system;
- *drawings*: topographic maps representing the territory features and, eventually, its height behaviour (contour lines and spot height spot);
- *images*: having the geometrical properties of maps (georeferenced images).

The *numerical vector products* are based on the type of the measured objects and the way in which the coordinates are organized:

- *topographic maps*: 2D or planimetric maps where height information is supplied through contour lines and height spots. In these maps the coordinates, defining the vertices of the geometric primitives representing the objects, are bi-dimensional;
- *thematic maps:* they give information about specific territorial themes like road conditions, hydrography, land cover, slope and exposition.
- *height profiles*: they represent the way the terrain height varies along a line (profile) defined through the area of interest;
- *3D models of objects*: the geometric primitives are represented by their vertices coordinates, measured in a 3D mode; object volumes are inferable in this representation.

Different image products can be derived from original images through more or less complex geometric transformations:

- *warped images:* warping is suitable only for flat regions or objects where the altimetric movements are small enough to be disregarded at the representation scale. The relationship between the image and the object spaces is approximated with flat transformations;
- *orthoimages: orthoprojection* is an approach that applies the rigorous 3D geometrical model relating the object and the image spaces taking care of the vertical movements of the object space. This allows the correction of positioning errors resulting from the recoding system or from the object morphological irregularity (not flat surface)
- *image mosaic:* it results from the union of adjacent single previously warped or orthoprojected images; it represents a wide image covering the whole area of interest.

According to the data representation, it is possible to distinguish:

- *analogical photogrammetry:* if the image is generated and supplied to the process on a photographic support (film, slides or printings);
- *digital photogrammetry:* if the image is supplied and processed in digital format. Digital image can derive directly from digital acquiring systems (sensors) or from

digitization, by desktop scanners, of hardcopy format images (traditional ones). In this sense digital photogrammetry is near to remote sensing.

According to the recording instruments operating in a digital mode, the following can be distinguished:

- images where the geometrical relationship linking the image and the object spaces is exactly the central perspective (frame acquisition);
- images where the geometrical relationship linking the image and the object spaces is something more complex than the traditional central perspective, generally a multiple central perspective (pushbroom and whiskbroom scanning systems).

Even if this second kind of image is becoming more and more diffused, mainly due to high-resolution satellite missions, traditional photogrammetry has always considered images generated through a single central perspective geometry (Fig. 3.4). These images are called *frame image* and the approach to process them is called *traditional*.

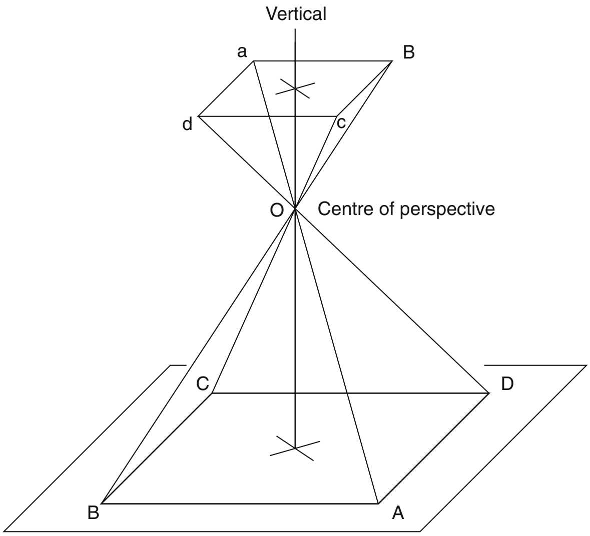


Fig. 3.4 The central perspective geometry linking the object and the image space

3.4 Traditional Photogrammetry

In frame images (both digital and analogical), objects are represented through a rigorous central perspective: the centre of projection is coincident with the focal point of the acquiring optical system. Traditional photogrammetry is

- *terrestrial* if images refer to objects located on the Earth's surface, or next to it. It is carried out using photo cameras positioned at ground level (e.g. building surveys, landslides monitoring);
- *aerial* or aerophotogrammetry, if the acquisition is carried out from above. In this case, the photo camera is on board aircraft and the object is the ground as it appears from that higher position. Present maps are derived from the application of this kind of survey.

The photogrammetric process can be resumed as follows:

- image acquisition/recording;
- image orientation and stereo-model generation;
- stereo-plotting that is object measurement inside the generated stereo-model;
- eventual orthoimage generation.

3.4.1 Stereoscopy and Restitution

The third dimension of the objects, i.e. the sense of depth intended as reliefs or depressions of the observed object, can be obtained only through a binocular vision. Eyes, observing the same object at the same time from two different points of view (*convergence*), induce the brain to join the two images, generating a 3D vision. In photogrammetry, the acquisition of the same scene from two different points of view, or centres of projection, enables, under certain conditions, a 3D vision, called stereoscopic, of the observed object to be obtained (Fig. 3.5).

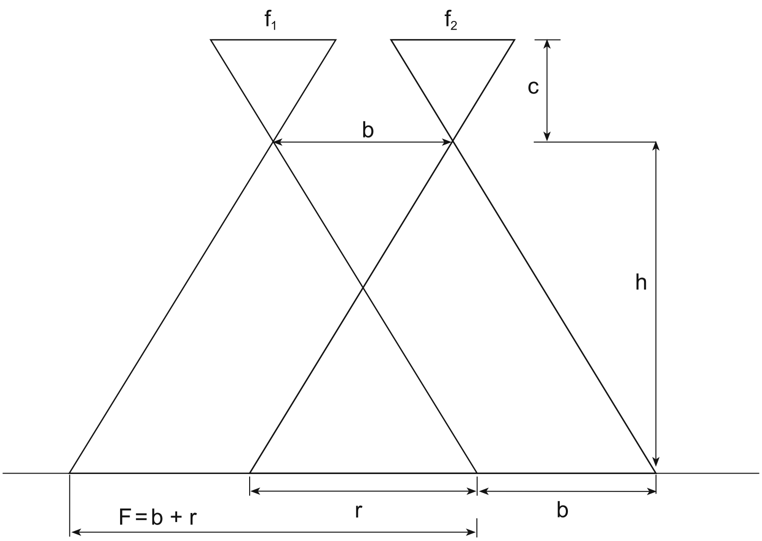


Fig. 3.5 Scheme of stereoscopic acquisition of aerial photogram and/or images

The following conditions must be respected to allow stereoscopy:

- images used for stereoscopy must at least partially cover the same area; the stereoscopic effect will be possible only over the overlapping part;
- acquisition must be carried out at a constant distance from the object, so that all the photograms have approximately the same scale;
- if the acquisition is aerial, requiring more than a single strip (see Chapter 6), strip direction must be as parallel as possible;

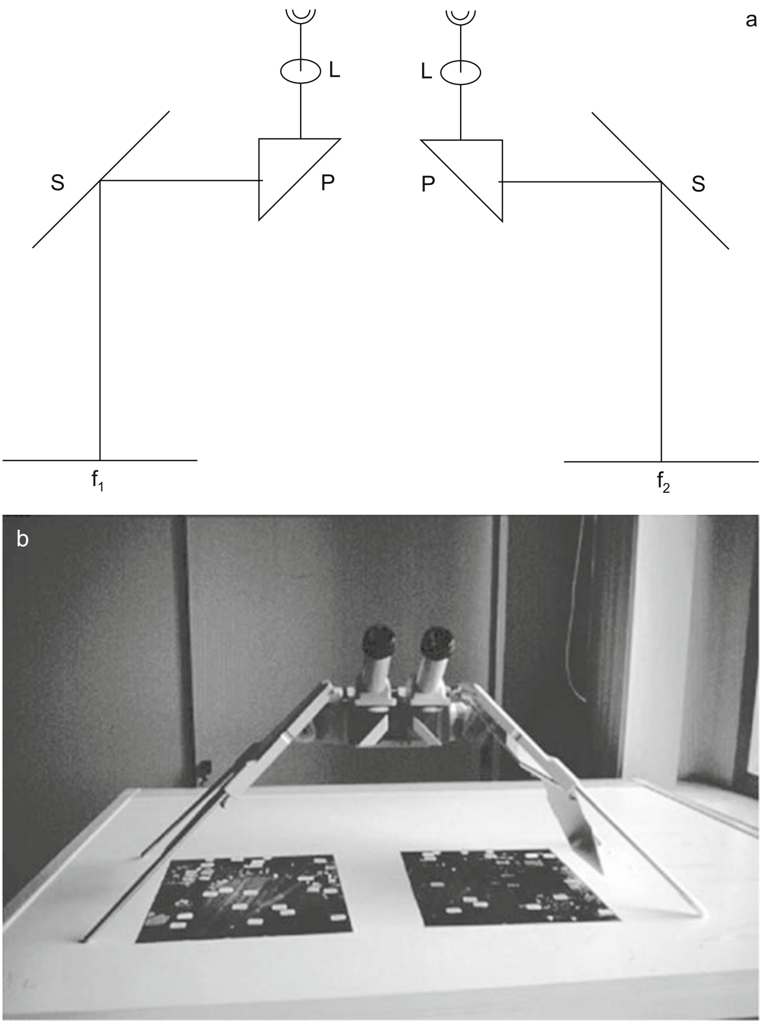
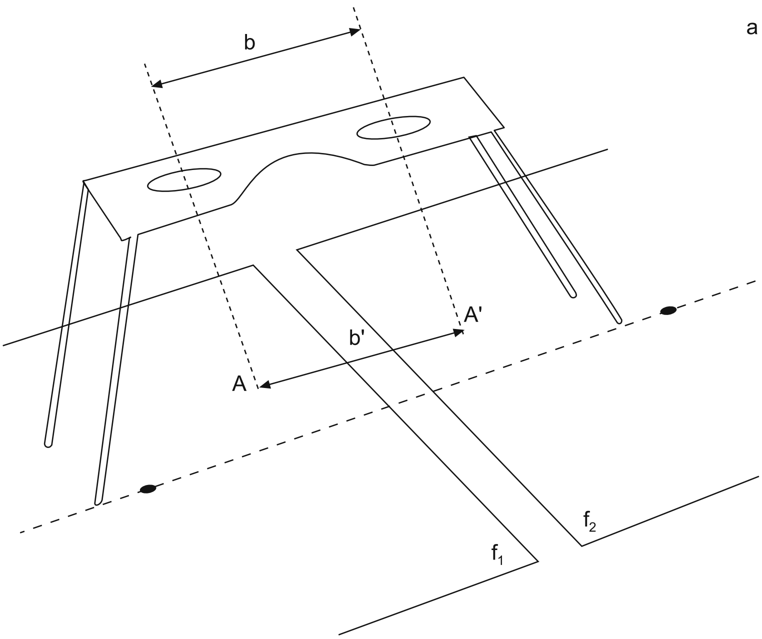


Fig. 3.6 (**a**) Conceptual system of a stereoscope with mirrors and prisms and (**b**) table stereoscope

- minimum distance between the acquisition centres of the two photograms participating in the restitution must be respected: in particular, the ratio between the distance of two subsequent acquisition centres (baseline B) and the distance between the observed object and the line joining the two acquisition centres (Z) usually have a basic ratio of 1/2, 1/3 (Figs. 3.6 and 3.7).



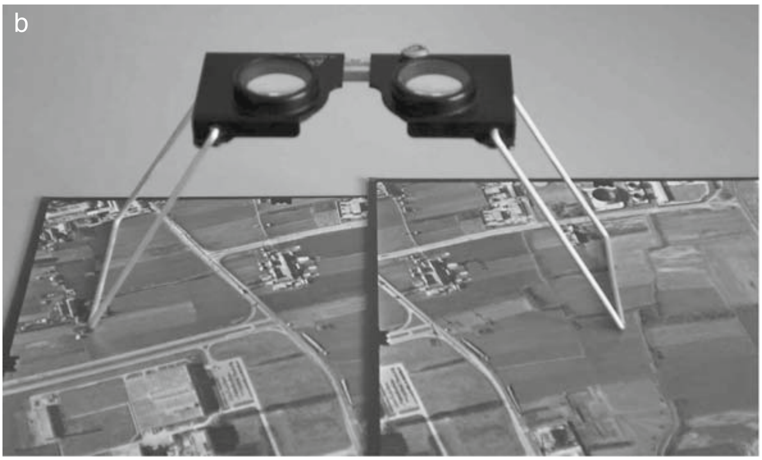


Fig. 3.7 Viewing of photographic stereo-pairs 23 cm × 23 cm with a pocket stereoscope. (**b**) Inter pupillary distance and (**b**′) stereoscopic base

In operative terms stereoscopy between two photograms is produced only through controlling horizontal and vertical *parallaxes*. Before defining parallaxes, the concepts of homologous points and homologous rays must be introduced:

- *homologous points*: points representing the same object in different images;
- *homologous rays*: projective rays corresponding to the homologous points joining the object-point with its correspondent image point.

The *horizontal parallax* (dx), or linear parallax, lying on the visual plane containing the baseline and the observed object-point, is the so-called depth of the object. Through its measure the observed altimetric coordinate of the points, or relative depth, can be determined. Intuitively it can be imagined as a variation of the relative distance between two object-points located at different distances from the centre of acquisition, as a consequence of the shifting of this one. A typical case of horizontal parallax is when from the window of a moving car a near object is seen in a progressive relative movement in respect of a further reference point (Fig. 3.8).

The horizontal parallax can be noticed only for values of d $\delta\gamma = \gamma_{\Pi} - \gamma_{\Theta}$ (*stereoscopic acuity*, see Fig. 3.9) higher than about 3 mgon. For angles dγ higher than 1.3 gon (in the centesimal system the centesimal degree, or gon, is defined as 1/400 of the 360° angle) the human eye cannot join the two images and consequently stereoscopy is lost.

The vertical parallax (dy), or height parallax, is geometrically defined as missed intersection between two homologous non-complanar (which then do not intersect, i.e. are crooked) rays. The vertical parallax must be eliminated in order to allow the eye to see stereoscopically, as it can only perceive horizontal parallaxes.

In aerial photogrammetry, natural (i.e. without vertical parallaxes) stereoscopic vision conditions can be obtained, just in a normal condition, which is when both consequent frames are nadir (perpendicular to the ground and relatively parallel) (Fig. 3.9). Ordinary operative conditions always require vertical parallax processing before proceeding with stereo-plotting operations.

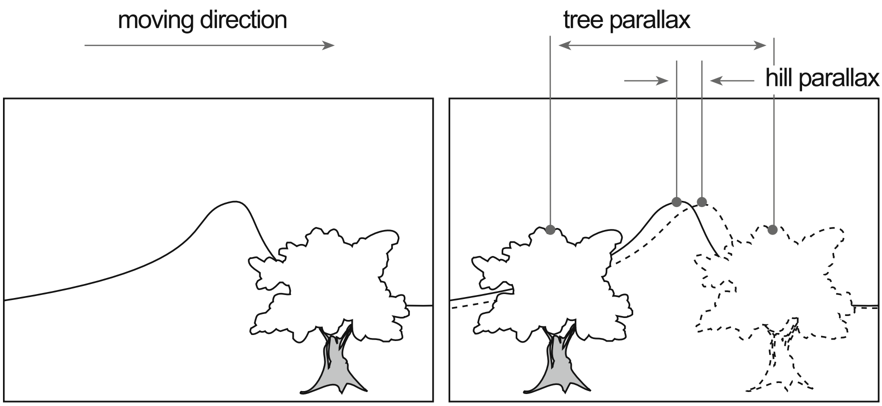


Fig. 3.8 Intuitive concept of horizontal parallax resulting from the sequential acquisition of the same scenes from different points of view

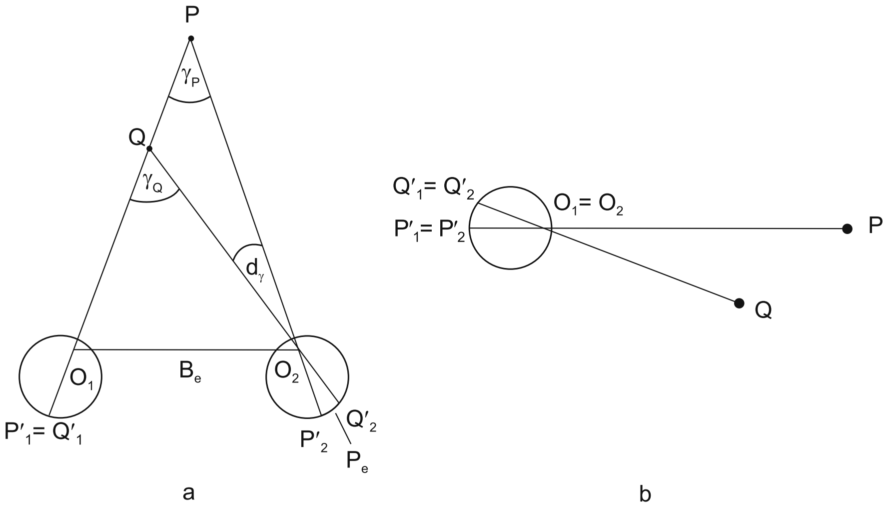


Fig. 3.9 Natural 3D view and parallax measures of images

3.4.2 Geometrical Basics of Photogrammetry

Historically photogrammetry has always considered the central perspective as its operative foundation. This approach results from photographic cameras' acquisition procedures: they produce photograms, whose main peculiarity is to be instantly generated through a geometric scheme unique for the whole image. This means that each object-point can be rigorously related to its correspondent image point by correctly defining the geometrical relationship linking the two involved spaces (object and image). The geometrical relationship is the central perspective, which is a suitable model whose parameters are identical for the whole image. The model parameters estimation is one of the basic steps of the photogrammetric process, and it is known as image orientation.

The image space, as illustrated in the scheme of Fig. 3.10, is defined by its reference system having ξ and η axes and origin in the Principal Point (PP).

ξ_0 and η_0 coordinates of the Principal Point PP and the *focal length c* locate the position of the projection centre on the image plane and define the *internal orientation* of the photogram. As well as these parameters, the radial distortion affecting the faithfulness to reality of this geometric scheme must be considered. In modern aerial photogrammetric cameras, this effect can be disregarded as usually lower than the precision with which image points' coordinates can be measured.

The coordinates of the perspective centre (X_0, Y_0, Z_0), expressed with respect to the object reference system (generally the cartographic system, see Chapter 2) and the three rotation angles ω, ϕ, κ describing the photo camera position at the moment of acquisition define the *external orientation of the photogram* (Fig. 3.11).

For the rigorous description of the acquisition geometry of the image, *nine parameters* have to be known. The internal orientation parameters (3), provided by the constructor, are assumed constant and specific for the photo camera.

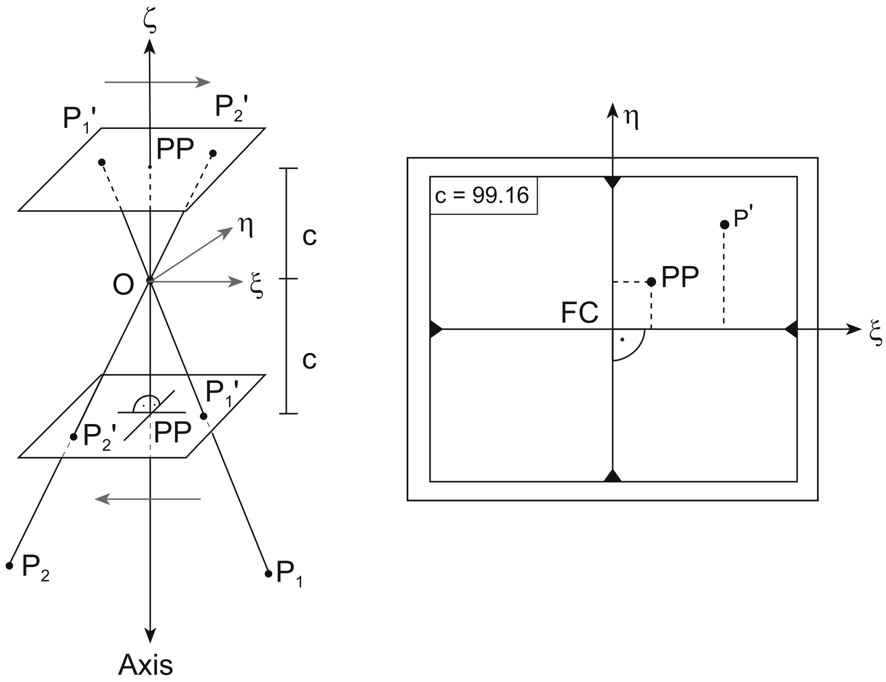


Fig. 3.10 Internal orientation: characteristic points and adopted conventions: O: Centre of projection, or point of acquisition; PP: Principal point with coordinates ξ_0 η_0 ζ_0c: focal length or constant of the photo cameraFC: Fiducial Centre, intersection between the lines that connect the opposite Fiducial Marks

The external orientation parameters (6) can be derived:

a) From position (GPS) and attitude (IMU, *Inertial Measurement Unit*) measurements carried out during the acquisition: in this case the approach is called direct *georeferencing*.
b) More commonly, through a Least Squares estimation process performed with respect to the Ground Control Points (GCPs).

Collinearity equations formalize direct (backward) (3.1) and inverse (forward) (3.2) relations among the object coordinates (X,Y,Z). They obviously depend on the internal and external orientation parameters.

$$\begin{aligned}\xi &= \xi_0 - c\frac{r_{11}(X-X_0)+r_{21}(Y-Y_0)+r_{31}(Z-Z_0)}{r_{13}(X-X_0)+r_{23}(Y-Y_0)+r_{33}(Z-Z_0)}\\ \eta &= \eta_0 - c\frac{r_{12}(X-X_0)+r_{22}(Y-Y_0)+r_{32}(Z-Z_0)}{r_{13}(X-X_0)+r_{23}(Y-Y_0)+r_{33}(Z-Z_0)}\end{aligned} \tag{3.1}$$

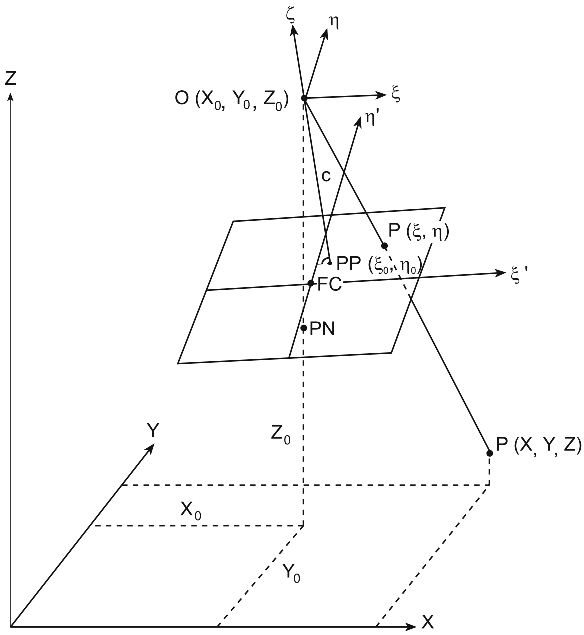


Fig. 3.11 Relationship between the coordinates of image points and object-points (Kraus, 1998). ξ_0. η_0. ζ_0.: image coordinates x,y,z: coordinates in a local system X,Y,Z: coordinates in the object system, or ground coordinates

$$\begin{aligned} X &= X_0 + (Z - Z_0)\frac{r_{11}(\xi - \xi_0) + r_{12}(\eta - \eta_0) - r_{13}c}{r_{31}(\xi - \xi_0) + r_{32}(\eta - \eta_0) - r_{33}c} \\ Y &= Y_0 + (Z - Z_0)\frac{r_{21}(\xi - \xi_0) + r_{22}(\eta - \eta_0) - r_{23}c}{r_{31}(\xi - \xi_0) + r_{32}(\eta - \eta_0) - r_{33}c} \end{aligned} \tag{3.2}$$

Attitude parameters (ω, ϕ, κ) are composed inside the r_{ik} coefficients, which represent the elements of the spatial rotation matrix R defined as follows:

$$R_{\omega\phi\kappa} = \begin{pmatrix} \cos\phi\cos\kappa & -\cos\phi\sin\kappa & \sin\phi \\ \cos\omega\sin\kappa + \sin\omega\sin\phi\cos\kappa & \cos\omega\cos\kappa - \sin\omega\sin\phi\sin\kappa & -\sin\omega\cos\phi \\ \sin\omega\sin\kappa - \cos\omega\sin\phi\cos\kappa & \sin\omega\cos\kappa + \cos\omega\sin\phi\sin\kappa & -\cos\omega\cos\phi \end{pmatrix} \tag{3.3}$$

Anticlockwise rotation ω,ϕ,κ respectively around the axes X, Y, Z are to be intended as sequential; they are defined as follows:

ω: primary rotation, or transversal;
ϕ: secondary rotation, or longitudinal;
κ: tertiary rotation, or skidding.

Eq. (3.1) shows that each object-point is corresponded by one image point, while Eq. (3.2) shows that for each image point infinite object-points depending on Z exist. Thus using one only photogram it is impossible to model the 3D geometry of an object represented in the photogram.

While approaching the problem through a Least Squares Estimation not only the six E.O. parameters can be estimated, but, if unknown, also the three (or more if even radial distortion is taken into account) I.O. parameters can be estimated on the basis of a sufficient number of GCPs: in this case the approach is called self-calibration. This section omits to treat this object, described in Kraus (1998).

3.4.3 The Real Model: Distortion and Calibration

The application of the central perspective theoretical model to the real case must deal with approximations which oblige the adoption of appropriate solutions to minimize them (that appear as wrong relative positioning of image points in the internal reference system). In particular, the following must be considered:

- photo camera errors;
- lens system errors;
- photogram errors.

While photo camera errors can be referred to as constructors' errors, lens system errors are mainly related to imperfect co-axiality (with respect to the optical axis) of the centres of all the spherical surfaces of the lenses constituting the lens, to the imperfect assembling of the lens cone on the photo camera's frame, and/or to the lenses' deforming geometry.

Calibration medium curves (of radial and sometimes tangential distortion) defined in the laboratory through optical goniometers allow correction of the real geometric model, relating it to the theoretical one and allowing residual distortions lower than 5 μm (lens cones practically without distortions). This information is reported on the warrantee certificate of the photo camera, which is then an essential instrument for the exact reconstruction of the photogram internal orientation.

Photogram errors are to be attributed to the support deformations, film's stretching, four fiducial marks collimating (Fig. 3.12).

3.4.4 Instruments and Modality of Acquisition

Fundamental elements of the aerial acquisition are
- photo camera (analogical aerial photo camera);
- photogram;
- flight characteristics.

Digital cameras are described in Chapter 6 (Table 3.1).

Photogrammetric cameras used in aerial photogrammetry, due to their geometric characteristics and flights high costs, are generally of metric type also whereas the interpretative demands prevail in comparison to the metric ones (Box 3.3).

Fig. 3.12 Aerial photogram, original dimension 23 cm × 23 cm with evidence of the flight parameters, the fiducial marks in the corners and the photogram number

Table 3.1 Classification of the aerophotogrammetric camera based on their Internal Orientation (I.O.)

Camera	Characteristic
Metrical	I.O. known and constant in the time
Semi-metrical	With devices rebuilding I.O.
Not-metrical (Amateur)	I.O. unknown and variable in the time

The photo camera comprises the following elements:

- *lens cone*: optical system able to canalize luminous radiation so that the image point is correctly formed on the focal plane. Lens cones differ in focal length, that is the ability to focalize an object at a short distance with a wide field of view, i.e. wide-angle lens, or at a long distance with a narrow field of view, i.e. telephoto lens;

Box 3.3 Fundamental data reported on the calibration certificate of the photogrammetric camera. PPA, PBS and FC: characteristic points determined during the photo camera calibration and lying on the image plane

Image coordinates of the Fiducial Marks (small targets on the body of metric cameras)
Image coordinates of Principal Point of Autocollimation (PPA), Point of Best Symmetry, (PBS) and Fiducial Centre (FC)
Principal distance c
Curve of mean radial distortion
Calibration date
Information concerning the image resolution

- *shutter:* controls the film's exposure time;
- *aperture setting:* controls the aperture diameter allowing luminous radiation to enter the photo camera and to reach the film on the focal plane; the larger the aperture diameter, the higher the flux of light;
- *exposure metre*: measures the incident radiation intensity and defines the shutter speed in relation to the aperture setting to control the exposure; an excess of illumination causes over-exposure and under-exposure;
- *film magazine*: the mechanism to advance and metre the film and to flatten the film during exposure.

The process of acquisition of a photographic image steps is as follows:

- the photo camera lens canalizes the light so that every object-point in the reality corresponds to a punctiform image on the film;
- the shutter aperture controls the quantity of light entering the photo camera;
- the shutter release time defines the exposure duration;
- the image is recorded on the film.

In the particular case of the aerophotogrammetric camera, a central shutter ensures that the external and internal orientations are constant in the whole photogram (a condition that is not verified in scanning systems). Due to the particular conditions when a camera is used on board an aircraft, special solutions are needed, like radio interference suppression, 28 V direct electric power, conforming to flight techniques setting.

A photogrammetric acquisition system, made up of materials having similar thermal expansion coefficients, must also have (Fig. 3.9):

- a levelling system to keep the system horizontal;
- a control and drift correction system connected with the camera support;
- an interface to communicate with an efficient navigation system;
- a longitudinal overlap regulator able to keep the longitudinal overlap constant, also in the case of strong variations of height;

- a Forward Motion Compensation (FMC) device, which compensates for the aircraft movement to ensure the apparent immobility between the film and images which move on it;
- a system synchronously exposing the fiducial marks and the auxiliary data (Fig. 3.13) with the main shutter;
- a pneumatic system flattening the film over the frame.

Each photogram acquisition cycle, lasting about 2 s, is activated by an operator's impulse, for single photograms, by an impulse coming from the navigation system, for a swath. During the acquisition, the operator carries on the shot procedures using a telescope provided with a special pointing system able to control drift and overlap (cinederivometer). *Drift* is the phenomenon for which the aircraft's real route does not coincide with the apparent one, due to a component of transversal motion, for example, caused by the wind.

Photogrammetry is required for many diverse applications and acquisition instruments and materials must be adaptable. This way, photogrammetric cameras may be

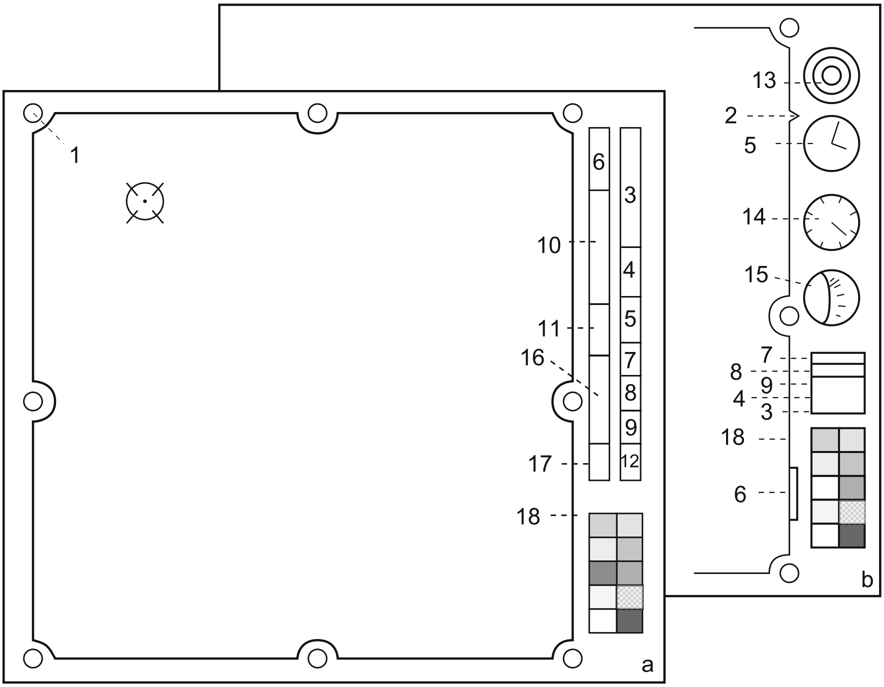


Fig. 3.13 Auxiliary data: (**a**) numerical and (**b**) analogical, reporting the photogram operational conditions of acquisition (Kraus, 1998) ***1*** *8 fiducial marks*, ***2*** 9th *asymmetric fiducial mark*, ***3*** *short project name*, ***4*** *date*, ***5*** *time*, ***6*** *photogram number*, ***7*** *camera number*, ***8*** *principal distance*, ***9*** *magazine number*, ***10*** *data for external orientation*, ***11*** *overlap*, ***12*** *photographic scale*, ***13*** *spherical equalizer (level)*, ***14*** *altimeter*, ***15*** *precision altimeter*, ***16*** *exposure time, lens aperture*, ***17*** *forward motion (length)*, ***18*** *grey wedge*

equipped with lens cones having different fields of view, with interchangeable filters, or use films with different sensitivity.

Moreover, all the conditions about acquisition (auxiliary data) need to be registered on the photogram in order to work properly later.

The technical scheme of a traditional aerial photogrammetric camera shows that it resembles more a modular system than a unitary device (Figs. 3.14 and 3.15).

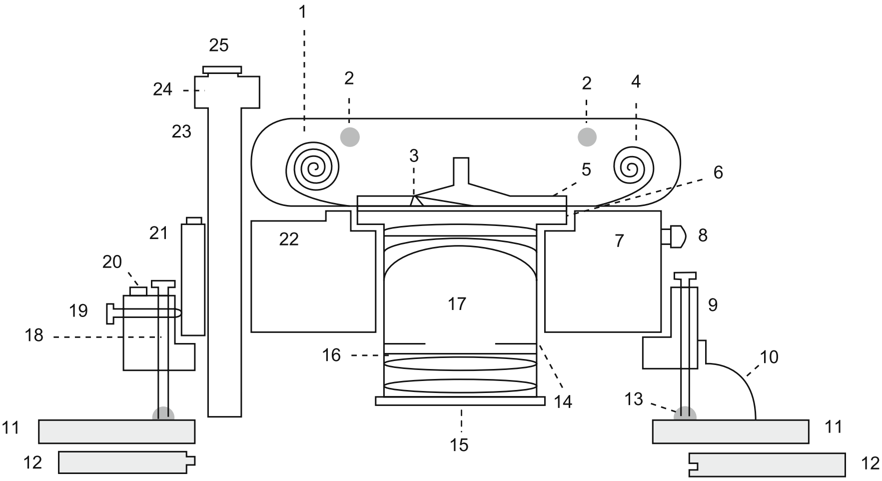


Fig. 3.14 Schematic diagram of a traditional aerial photogrammetric camera (Kraus, 1998) **1** photographic film magazine, **2** film indicators, **3** aspiration, **4** rollfilm width 23 cm, **5** pneumatic flattening system, **6** lens cone and camera frame, **7** electronics monitoring, **8** 28 volt CC, **9** camera support, **10** cable mass, **11** airplane bottom, **12** door, **13** supports anti-vibration, **14** lens aperture, **15** filter, **16** shutter, **17** lens, **18** 3 footscrews for ϕ and ω, **19** orientation K (K-clamp), **20** spherical level, **21** control instruments and switches, **22** transport control and aspiration, **23** viewfinder telescope, **24** cinederivometer, **25** ocular

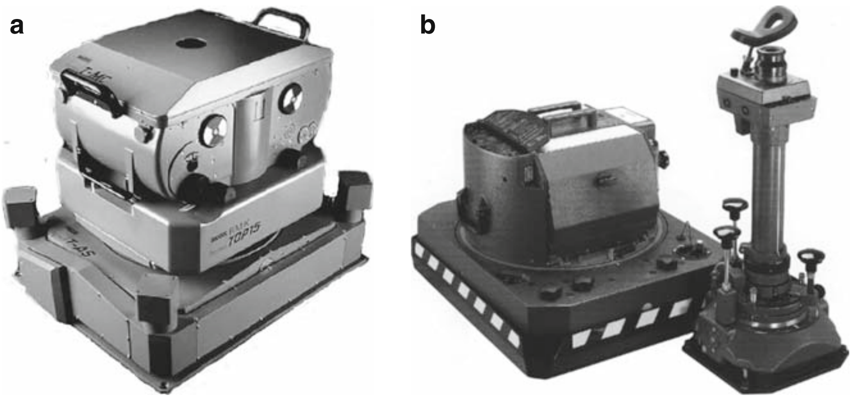


Fig. 3.15 Two among the several aerial photogrammetric cameras available: (**a**) *Zeiss aerial camera RMK TOP* (Terminal Operated) and (**b**) Leica aerial camera RC30 Wild

3.4.5 Flight Plan

Before performing an aerial photogrammetric acquisition, a flight plan need to be designed taking into account operative variables. To do this many geometric relations must be considered (Fig. 3.16).

A flight is planned as adjacent, generally parallel, strips whose relative position depends on the acquisition systems, on the requirements of the stereo-plotting step and on the accuracy and typology of the final product.

The strip is a sequence of photograms acquired along the same flight direction.

In general the trend is to fly at the highest altitude possible as regards the accuracy and image-interpretation conditions, so that the area of interest can be covered with the smallest number of stereoscopic models in order to reduce plotting time and costs.

The flight path must be defined so that each strip is as long as possible in order to reduce aircraft turn and alignment time. Longitudinally adjacent photograms (along a strip) must share an overlap of 60–70%, transversally adjacent photograms

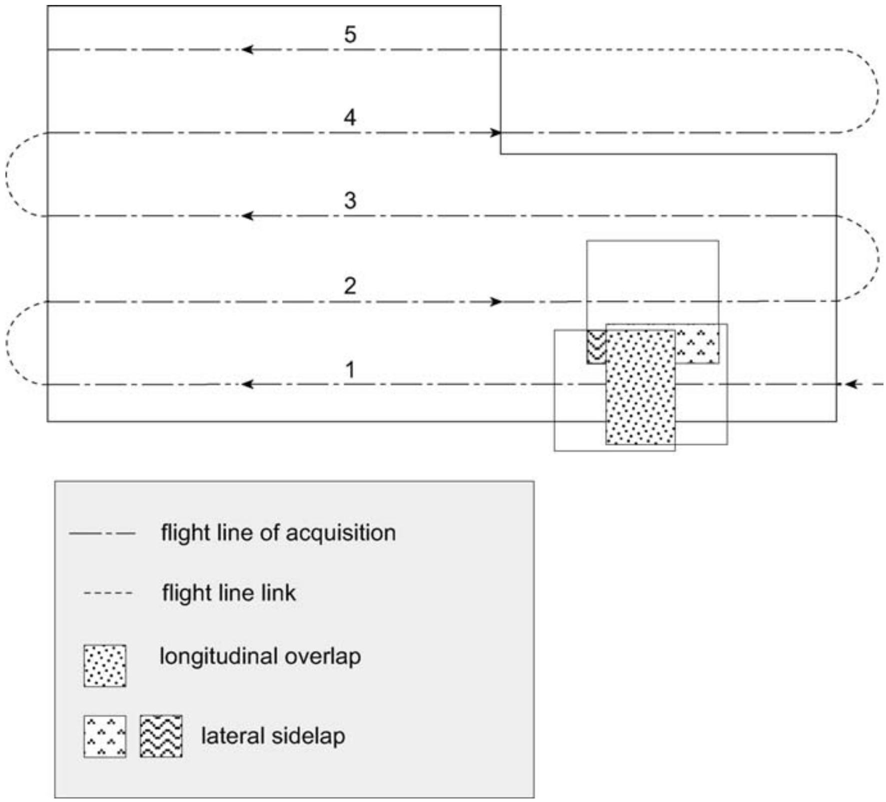


Fig. 3.16 Flight plan: geometric acquisition over a flat area

(between strips) must share an overlap of 20–30%. In this way the stereoscopy is assured everywhere.

The Flight Plan is intended as an integrated document containing all the details about the acquisition and the *flight path plan*.

With reference to Fig. 3.11, some geometric parameters conditioning the flight plan draw are reported in Box 3.4.

Box 3.4 Geometric parameters to be used preparing the flight map (see Fig. 3.16)

Scale of the photogram	$m_b = h/c$
Ground side of the photogram	$S = s \times m_b$
Base of the photogram	$b = B/m_b$
Relative flight altitude	$h = c \times m_b$
Absolute flight altitude	$Z_0 = h+Z$
Swath length (%) A	$l = 100 \times (S - B)/S = 100 \times (1-B/S)$
Swath width (%)	$q = 100 \times (S-A)/S = 100 \times (1-A/S)$
Covered area of the photogram	$F_b = S^2 = s^2 \times m_b{}^2$
Base for the overlap of photograms (%)	$B = S \times (1-l/100)$
Interaxis between swaths for trasversal overlap (%)	$A = S \times (1-q/100)$
Number of models per strip	$n_m = [1+L/B]$
Number of photograms per strip	$n_b = n_m+1$
Number of swaths per block	$n_s = [1+Q/A]$
Area covered by a model	$F_n = (S-B) \times S$
Increase of useful area per model	$F_n = A \times B$
Shutter release (s)	$\Delta t = (B\ [m] / v \times [m/s]) \geq 2.0$

where c: focal length; Q: block width; s: image side at the border; Z: medium ground altitude; L: swath and block length; ν: airplane ground speed (m/s)

The flight plan must be communicated as the scheme of the aircraft paths flight height. It must be overlaid on a topographic map and it must contain

- strip-oriented axes with the indication of the beginning and the end of the stereoscopic cover;
- absolute flight height for each strip;
- any obstacle to the flight;
- areas with restrictions over which flying is forbidden.

Auxiliary data about materials, instruments to be used, generic conditions must complete the flight plan.

3.4.6 *Artificial Stereoscopy Techniques*

Stereoscopy is realized in different ways, dependent on the format of processed images (films or digital images); operative conditions to be respected are common to all of the techniques and refer to the following criteria:

- horizontal parallaxes must be parallel to the ocular base;
- vertical parallaxes must be absent;
- convergence angles must not be higher than 1.3 gon;
- the difference in size of homologous objects appearing in two photograms (scale difference) must not be over 14%;
- the difference of photographic density (grey tones bar charts) should not be too high.

As regards the two first points, there would be no problem if the acquisitions were normal, while, in practice, the acquisitions are pseudo-nadiral and thus it is necessary to correct vertical parallaxes. As the correction has local validity, stereoscopy for pseudo-nadiral photograms can be obtained only for limited portions.

- The *stereoscope* is a simple optical instrument made of two lenses allowing the independent observation of two photograms giving a 3D vision of the object. In a lens stereoscope, stereoscopy is realized by putting a pair of images acquired in normal set under the eyes and observing them with a convergence equal to zero, i.e. with parallel ocular axes. In this condition, homologous rays of the object-points in common between the two photograms intersect the photograms plane, giving a 3D image. A pair of photograms is put in the stereoscope and, with proper planimetric movements of a photogram with respect to the other, stereoscopic vision of the whole stereogram is obtained. They are used for quick recognition of photographic details and in analytical plotter (stereocomparators).
- The *anaglyphic technique*, applying only to greytone images, is based on chromatic complementarity and on the composition of two monocula visions. According to this technique, one photogram, usually the right one, is visualized in red chromatic scale, while the other, the left one, in cyan chromatic scale (blue + green), which is the complementary of red. The observer, using a pair of glasses with a cyan-coloured right filter and a red-coloured left filter, ensures complementarity with respect to the images. This way, the left eye sees, in greytones, the left image, but not the right image, likewise the right eye only sees the right image. Optical fusion generates a greytone 3D image due to the adopted colours' complementarity. This method is used in optical projection analogical plotters.
- The *polarization technique* is a stereoscopic visualization system based on the two images observed through a polarizing filter, alternately according to the two perpendicular planes, and the use of polarizing glasses. This way each eye observes just one of the two photograms and generates the third dimension through the fusion. This technique is also suitable for colour images.
- *Liquid crystal shutters* are flat planes containing a film of liquid crystals becoming transparent or opaque according to the received impulses. The two images are sent to the screen at alternate frequencies; the shutters are set on glasses and are activated with (distinct) frequencies synchronized with the images projected on the screen's frequencies. The two eyes receive their respective separated images at alternate instants, with such high frequencies as to be considered as simultaneous reception. The glasses are defined as active as they receive impulses and

they work as a consequence of this; synchronism with image visualization on the screen is realized by an infrared cell (emitting) above the screen.

During the stereo-plotting operation, a 3D deformation of the image occurs, mostly regarding an exaggerated perception of altitude differences which also helps the operator to detect shallow differences in height.

When observing the stereoscopic model, the ratio base/apparent distance (stereoscopic observation distance) B_e/D_e is different to the ratio baseline/flight height, B/H, and it is usually $B_e/D_e > B/H$.

Thus in the stereoscopic model an exaggeration of altitudes e_z occurs, given by the ratio:

$$e_z = \frac{B}{H} : \frac{B_e}{D_e} \tag{3.4}$$

where:

B: baseline
Be: base of stereoscopic observation or interpupillar distance (about 65 mm)
H: flight height
D_e: apparent distance of the image from the observer (in stereoscopes it is 35–60 cm).

3.4.7 Image Orientation and Stereo-plotting

To perform the photogrammetric projection, the acquisition geometry must be modelled; thus it is necessary to properly orient the photograms in order to eliminate vertical parallaxes, resolve the horizontal ones and correctly formulate collinearity equations. The problem can be conceptually faced in different ways. The adopted technique and instruments condition the choice.

The operation enables on the one hand the stereoscopic vision, hence the 3D observation of the object, and on the other hand the measurement of the observed points within the reference system, e.g. cartographic system.

In the ideal case, which applies to most satellites and aerial images nowadays, the model should be generated in a direct way, using photogram internal (c, $\xi_{0,}$ η_0) and external ($X_{0.}$ $Y_{0.}$ $Z_{0.}$ ω, ϕ, κ) orientation parameters, independently measured during the photogram acquisition. This is a revolutionary approach compared with the classic photogrammetric approach. Its real application potentiality is mainly related to the evolution of technical instruments and data processing in progress. The instruments playing the major role in collecting measures are satellite positioning systems (GPS, *Global Positioning System*) and integrated with the instruments of acquisition system attitude (inertial platforms INS/IMU, *Inertial Navigation System/Inertial Measurement Unit*) (C. Heipke, 2002), while the achieved calculation efficiency processes huge amounts of data.

Integrating Positioning System GPS registrations (see Chapter 7) with data collected through IMU instruments, in which a gyroscope's functions are combined with the functions of a high-precision digital accelerometer, allows the determination of both the single acquisition centres coordinates and the angular attitude of the camera optical axis, georeferencing the data.

Measures processing is part of the process that must deal with measured units variable in the time, whose variation is to be modelled in order to make a discrete sampling continuous and to manage possible anomalies by using strong data processing techniques (e.g. Kalman filter). These aspects of temporal variability, which might look secondary compared with classic photogrammetric acquisition (the photogram is generated at the same time), are very critical for digital acquisition by lines. Moreover the correct use of the measures requires managing subsequent roto-translations, which are necessary to pass from many local reference systems of the instruments to the object reference system.

Analogical photogrammetry is still based on orientation, requiring the use of GCPs. It is only through an accurate ground survey of well-defined points and their correct identification on the photograms that it is possible to estimate the external orientation parameters of images. Camera internal orientation parameters are reported on the calibration certificate released by the producer. The analytical solution to the problem of the estimation of the external orientation parameters can be achieved in different ways.

The development of analytical photogrammetry was granted by the development of the numerical calculus techniques that allowed optimizing of the limited memories of the computers of the 1970s for problems like solving big linear systems.

If a single stereo pair is considered, the possible approaches are

- *single image orientation* or pyramid vertex adjustment: the orientation of the two photograms of a stereo pair occurs independently using GCPs and without taking into consideration any relative relationship between the images of the pair. The suggested solution is analytical; the unknown parameters are the six parameters of single image external orientation. For this task, at least six equations of collinearity must be written, that is at least three plano-altimetric GCPs are required;
- *one-step stereo pair orientation*: method which counts to analytically determine the 12 unknown external orientation parameters relating to the two photograms (6+6) by solving a system of at least 12 equations of collinearity;
- *two-step stereo pair orientation* counts the formal separations of the equations, in order to previously orient a photogram with respect to the other one (relative orientation), and consequently the obtained stereo-model with respect to the object space (absolute orientation). It is what is conceptually performed with the analogical stereo-plotters.

The last method is also referred to as Independent Stereo-Model Adjustment. The solution is achieved exploiting a sufficient number of Tie Points (TPs) linking in a

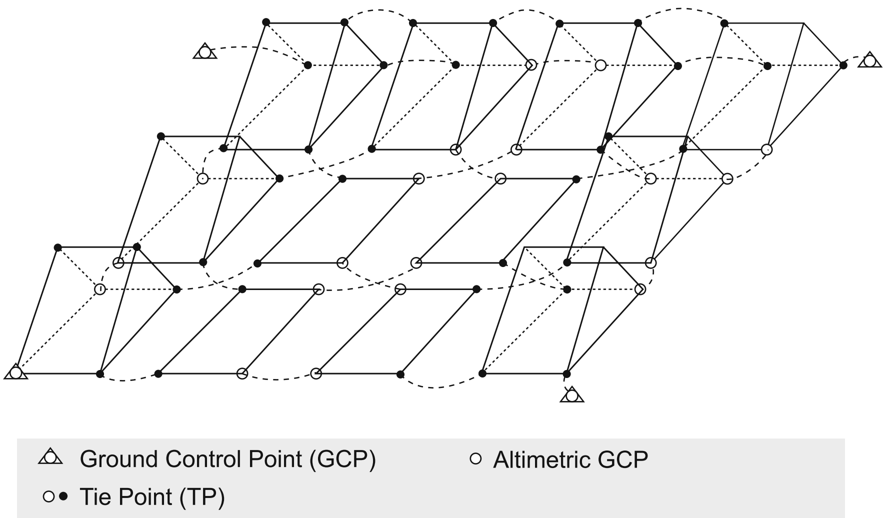


Fig. 3.17 Aerial triangulation or bundle adjustment. Conceptual scheme of compensation at independent models

relative way the two images of the model and a limited number of Ground Control Points (GCPs), well spread over the covered area (Fig. 3.17). GPSs are needed to correctly orientate the model with respect to the ground system through least squares compensation

Each approach requires the knowledge of a different number of GCPs that can be defined completely (3D coordinates) or partially (only planimetric or only altimetric).

In the specific case of a photogrammetric flight, in which orientation of many stereo-models is required in order to reduce the number of GCPs, Aerial Triangulation is adopted. This approach, also known as Bundle Adjustment, estimates at the same time the exterior orientation parameters of all the images of the whole block coming from the strips generated according to the already drawn flight plan (Fig. 3.18).

To simultaneously absolutely orientate (Aerial Triangulation) all the models of an aerial photogrammetric block, a sufficient number of solving collinearity equations must be declared. The Aerial Triangulation, also referred to as Bundle Adjustment or Projective Stars Adjustment, postulates the direct calculation of the relations between image coordinates and object coordinates for all of the images of the block, disregarding the intermediate step of the independent model orientation (a pair at a time). The equations of the solving system refer both to Tie Points and GCPs. TPs are needed to link in a relative way two images of the same stereo pair, or generally images of the block showing the same area. They are used in a Bundle Adjustment to increase the number of the required equations, in order to keep lower the number of GCPs (Fig. 3.19).

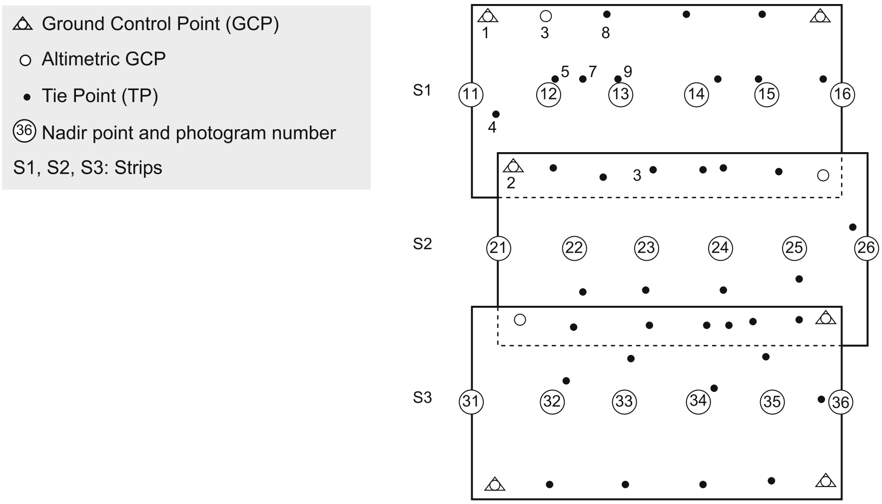


Fig. 3.18 Photogrammetric block with indication of *Ground Control Points (GCPs) and Tie Points (TPs)*

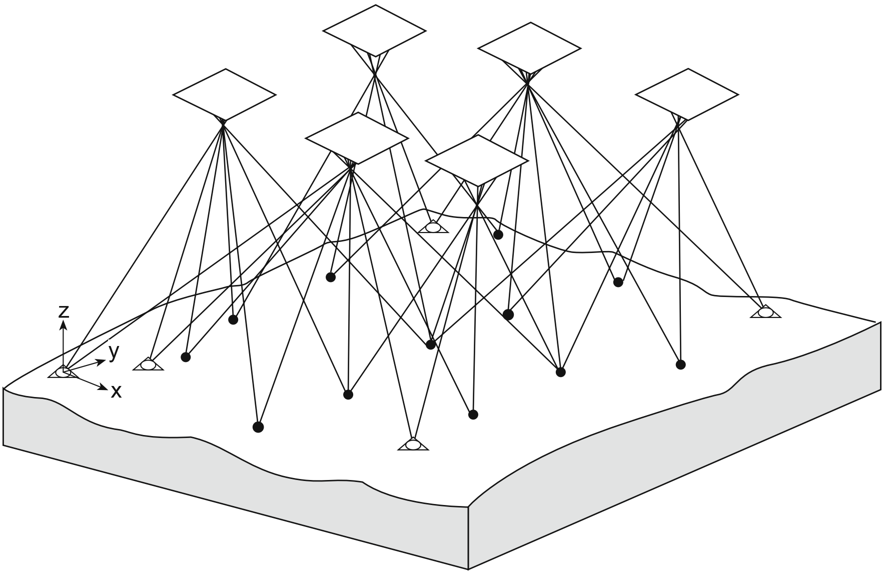


Fig. 3.19 Conceptual scheme of compensation at projective stars

During adjustment, adoption of planimetric points (2D) is suggested, along the block sides, and plano-altimetric points (3D), transversal to the flight direction, in order to increase the accuracy of the solution.

Ground control points acquisition by GPS instrumentation, compared with traditional topographic techniques, enables the survey in 3D, improving the precision of the results with advantages on costs (Box 3.5).

Box 3.5 Definition of ground control points, check points and tie points

Ground Control Points GCPs	Used for the estimation of the exterior orientation parameters of the selected sensor model. Residuals resulting from the Least Squares estimation measure the model accuracy, not the final product one
Check Points CPs	Used for the final product accuracy assessment. They do not take part to the exterior orientation parameters estimation process
Tie Points TPs	They are used while orienting a stereo pair or a stereo image block. In a Bundle Adjustment, they are used together with the GCP during the Least Squares estimation process, increasing the number of the declared equations In an Independent Models process, they are used to solve the relative orientation of a single stereo pair

3.4.7.1 Stereo-plotters

Stereo-plotting operation (sometimes referred to as restitution) is performed through stereo-plotters, which can be classified into three groups corresponding to the evolution of the photogrammetry:

- *analogical*, belonging to the history of photogrammetry until the present;
- *analytical*, universally employed, still often used. In these instruments the operator observes, as in a stereocomparator, two fixed collimation marks and photograms lying on two independent frames, and each can move in orthogonal directions along pairs of tracks. Commands to move the frames come from the computer and not from the operator. Some impulses are generated, counted in an electronic loop and transmitted to the computer in the form of increments of x, y instrument coordinates which faithfully represent, in the ratio 1:1, the image coordinates;
- *digital*, which are becoming more and more common and replacing the analytical ones. The restitution unit is an informatics system equipped with a device for stereoscopic vision (i.e. active glasses with polarized screen), an opportune hardware (pointers) for the stereo collimation, one or more calculation units for managing orientation and plotting. Therefore the stereo-model is video formed, and measures and processing are carried out thanks to specific software managed by powerful PCs.

3.5 Digital Photogrammetry

When the imagery processing steps, even including stereo-plotting, can be managed by a computer, photogrammetry is defined as digital. Suitable images can be

- photograms (slides), acquired by traditional photogrammetric systems then digitized with photogrammetric scanners;
- frame images with a single central perspective geometry generated by digital cameras;
- scanned images coming from digital systems with linear scanning (pushbroom), electronic or optical–mechanical (Plate 3.1).

In the first two cases, the adopted orientation and stereo-plotting techniques are the same as those used in traditional photogrammetry. Analytical procedures for image orientation are still managed by a computer which can also turn, in a single step, to stereoscopic vision and to 3D collimation of the video points. The stereo-plotting device is entirely represented by the computer and replaces the traditional analytic photogrammetric plotter.

If the image is generated by linear scanning digital systems, the issue of the acquisition geometry reconstruction passes through the adoption of more complex time-dependent projective techniques comparable to a multiple central perspective. This topic is covered in Chapter 6 on instruments for Earth observation.

3.5.1 Traditional and Digital Systems

The main difference between a traditional photographic camera and a digital camera is in the different way of registering an image (Fig. 3.20):

- film, in analogical cameras;
- by means of the dispositive *Charge-Coupled Device* (CCD), conceptually similar to the sensors used for satellite acquisitions and in digital cameras (Plate 3.1a and b).

Photographic acquisition in the traditional process consists of

- choice of photographic equipment: camera, lenses, film sensitivity;
- photochemical development and possible print;
- film scanning if the analogical acquisition is to be turned into digital;
- image digital processing.

The photographic process develops in subsequent operative and process steps, while with digital acquisition image availability and monitoring possibility are real time with immediate feedback on the acquisition's quality.

Except for the shot, or acquisition, processing and print operations take a long time, are expensive and complex. Whereas, in the case of digital images, the operator has total and immediate control of the operative phases.

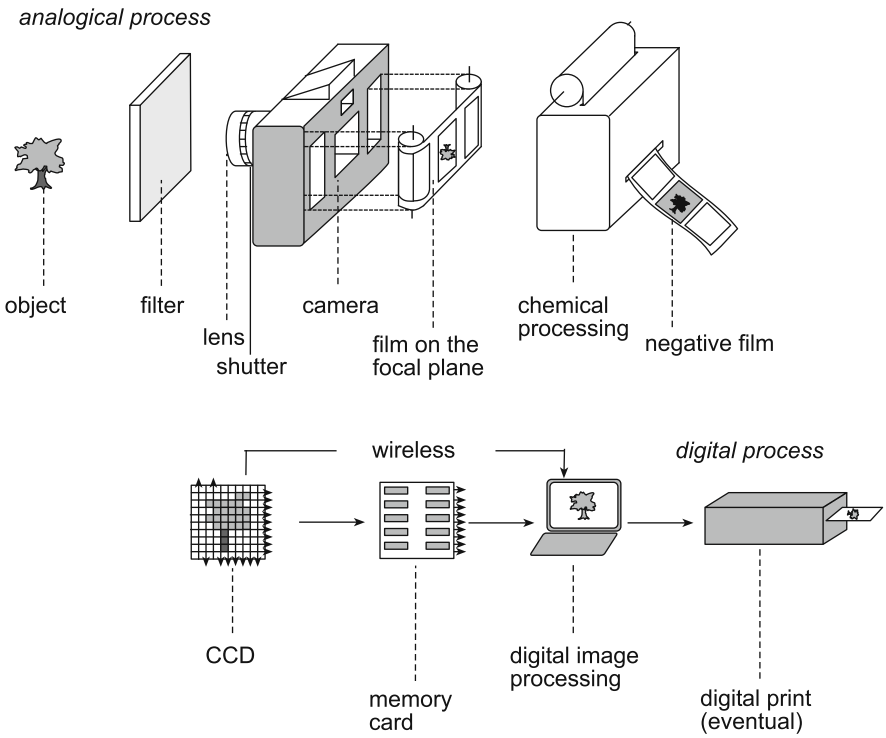


Fig. 3.20 Analogical (**a**) and digital (**b**) processes of image acquisition

Photographic images in grey tones are registered on an emulsion containing light-sensitive silver halide salt crystals. They react to the quantity of light reflected by the acquired scene.

Colour film, in addition to silver halide crystals, contains chromatic layers sensitive to yellow, magenta and cyan colours, complementary of blue, green and red. When a film is exposed to light the chemical structure of the crystals (crystals' size determines the grain) is modified, as the silver halide is transformed into metallic silver, which forms the image on the negative (see Plate 8.5).

In digital photography, the image is formed from an analogical electric signal which is transformed into a pixel by an analogical–digital converter through a sampling process. The pixels are assigned three different values of Digital Number (DN), corresponding to red (R), green (G) and blue (B). A processor's screen uses RGB values to reproduce the colour of each pixel.

The advantages of the digital in respect to the analogical photographic process can be summarized in the following points:

- high radiometric resolution;
- quicker availability of the image with the same acquisition time;

- digital processing accessibility;
- reproduction identical to the original, quickly and at low cost;
- possibility to be inserted in information systems and of management.

If an analogical image is scanned, hence transformed into digital format, the last three points subsist.

3.5.2 Format of Digital Images

Digital data storing formats have implications both about calculus and metric aspects.

A digital image is made by a numerical value-ordered sequence such as to allow the faithful reproduction of the (raster) image that it represents. Rules defining the data organization structure determine the digital image format.

These rules, in the specific case of photogrammetry, in which the quantity of data to process is huge, must respond to the following criteria:

- space saving;
- data compression and compression factor consistent with preserving the geometric integrity of the original image through reducing the occupied space;
- possibility to exhaustively represent the original radiometry;
- free use.

In the photogrammetric domain, two main formats are adopted: TIFF (Tag Image File Format) and JPEG (Joint Photographic Experts Group).

TIFF format characteristics are

- storing many images in one single file;
- adopting non-destructive compression and decompression system, Lempel Ziv Welch (LZW) able to halve the file size on RGB images.

JPEG format is a compressed format which can be generated in two ways:

- through non-destructive compression techniques, which preserves all the original information though reducing the file's physical dimensions;
- through destructive compression techniques, but with limited effects on the visible perception of the image (Discrete Cosine Transformation). In this case, data's partial degeneration appears as light out-of-focus effects in correspondence to the radiometric borders and as objects' positional movements on the image whose entity depends on the adopted compression ratio, making this technique not suitable for photogrammetric applications.

3.5.3 Digital Images' Metric Content

In the case of the digital image, the reconstruction of the image's physical dimension is not so immediate as in the case of the analogical in which the frame is a measurable physical object.

The digital image is a numerical matrix without physical dimensions which can be retrieved analytically.

In fact analytical relations allow the position of each element (pixel) of the matrix to be transformed into coordinate pairs of the element itself, provided that the scanning geometry from which the image comes is known. In particular the scanning geometric resolution must be known (i.e. the pixel dimensions Δx, Δy), whether it is of the aerial/satellite digital sensor or of the photogrammetric scanning system used to digitize a photogram.

The image reference system conventionally adopted is reported in Fig. 3.21.

The relationship linking the matrix reference system (column i, row j) to its physical correspondent (image system) is defined as follows:

$$\begin{aligned} X_{i(barycentre)} &= i \cdot \Delta x \qquad i = \text{int}\left(\frac{X_i}{\Delta x} + 0.5\right) \\ Y_{j(barycentre)} &= j \cdot \Delta y \qquad j = \text{int}\left(\frac{Y_j}{\Delta y} + 0.5\right) \end{aligned} \tag{3.5}$$

In the case of the digital image coordinates traditional measurement involves locating a cell through its matrix coordinates (i, j). The relations above allow translation of the matrix coordinates into image coordinates, referred to the cell barycentre.

Matrix cells' geometric resolution is commonly indicated by the sampling density, which defines the number of cells in the image contained in a certain length

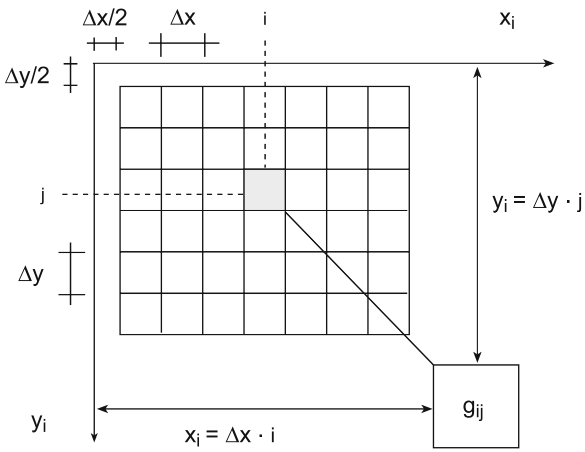


Fig. 3.21 Image reference system

unit. The most used parameter to indicate an image sampling density is the dpi (dots per inch), indicating the number of cells per inch (= 25.4 mm). In general the sampling densities in the two spatial directions coincide; this way the cell results in a finite size square.

When proceeding at photograms digitization, considerations about identifying the most appropriate resolution are needed. The choice is conditioned by the

- resolving power of the human eye (6–8 lines/mm, 300–400 dpi);
- information content of traditional photograms (80 lines/mm, 4000 dpi);
- screen's visualization capacity (about 100 dpi);
- memory occupation by the obtainable file.

The ideal solution would be to transport the geometric resolution of the analogical (4000 dpi) into the digital. This way the image's faithfulness would be guaranteed, although producing a huge amount of data to be managed. Actually the techniques used adopt a more rational approach based on digital processing and precision measures required by the final product.

3.6 Digital Photogrammetry Devices

Traditional photogrammetry needs complex optical–mechanical instrumentation (dial gauge and stereoscope) to proceed to photograms' orientation and restitution. In the case of the digital, instead, all functions are gathered within a computer where they are expressed through software. This transformation requires a certain level of complexity of the software itself and hardware's improved count skills, both for processing and visualization.

From the analytical point of view, orientation geometric models, at least in the case of digitized photograms, are nearly identical to the ones adopted in the analytical plotters.

Moreover data digital format looks at the experience gained in remotely sensed images processing in order to automate the process. In fact the restitution software can automatically detect, or collimate, fiducial marks, pre-pointed out points and homologous points, be they related to points defined by the operator or to points detected by the software itself based on opportunity criteria.

Regarding this analysis, the advantages and disadvantages of digital photogrammetry compared with traditional methods can be summarized as follows:

Advantages

- larger number of users: the informatics instrument is more familiar to an average operator;
- reduction of instruments' size and costs;
- reduction of work time due to the automation of some phases;

Disadvantages

- software designing is more expensive and complex;
- the required informatics instrumentation must have advanced processing characteristics.

Nevertheless, it is still necessary that acquisition (and/or photogram digitization), orientation and control phases are carried out by expert operators.

3.6.1 Digital Photogrammetric System

This system can be schematized in its two main components (Fig. 3.22):

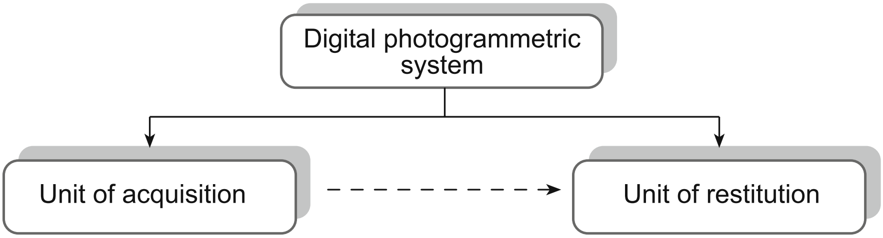


Fig. 3.22 Digital photogrammetric system

- *acquisition unit*: schematically represented in Fig. 3.23, it is the operative module for image acquisition and its rigorous geometric description: each cell must be able to be associated with fiducial system coordinates corrected from the systematic errors. The acquisition can be of two kinds:
  - direct, if the image is acquired by a digital camera (Fig. 3.24);
  - indirect, if the digital image derives from scanned traditional photograms.
- *restitution unit*: the operative module for data photogrammetric processing: orientation, triangulation, restitution, DEM generation, orthoprojections.

3.7 Digital Orthophoto

The aerial photogram is a central projection of a portion of land of which it is a qualitative and geometric information permanent archive. To acquire a metric value comparable to the cartographic one, the photogram must be transformed into an orthographic projection (Fig. 3.25).

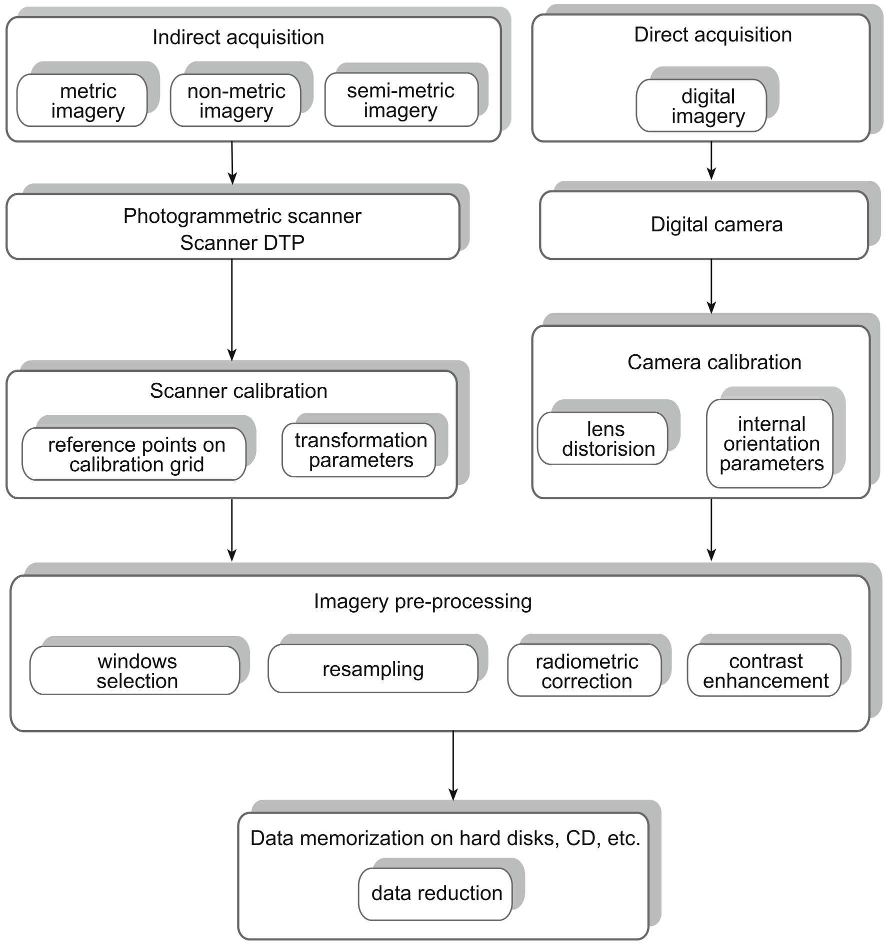


Fig. 3.23 Acquisition unit of a digital photogrammetric system; *DTP: DeskTop Publishing*

In the central projection of an aerial photo, each detail is in a planimetrically incorrect position in respect to the orthogonal projection of the corresponding Gauss map. This is down to two principal causes:

- photogram inclination in respect to the horizontal plane, or optic axis inclination in respect to the vertical;
- ground orography determining a continuous scale variation of the photogram in function of each considered point's relative altitude with respect to the acquisition point.

The digital orthoprojection is obtained by correcting each single pixel's position, which results shifted as effect of prospective deformations due to the altimetric

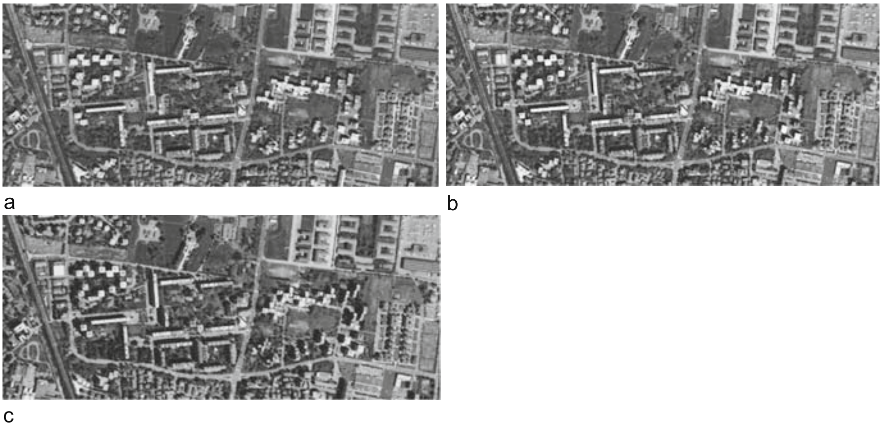


Fig. 3.24 Particular of a photogrammetric acquisition with a digital camera ADS40 in the three positions: (**a**) backward, (**b**) nadir and (**c**) forward (see Figs. 6.26 and 6.29). The different angles of acquisition are detectable observing the building displacement. These images are used in pairs for the stereoscopic view *(© CGR, Parma)*

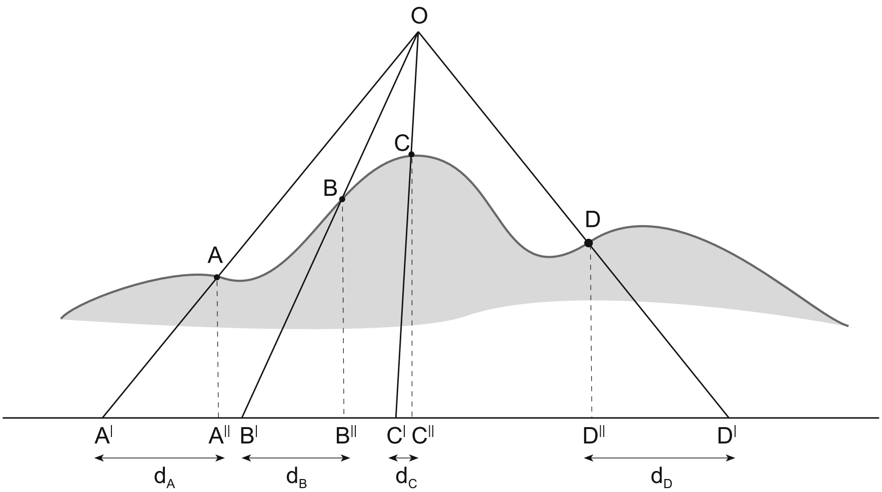


Fig. 3.25 Effects of the orography on the planimetric positioning of the points A, B, C, D. dA, dB, dC, dD: errors of planimetric positioning

variations described in detail by a properly georeferenced Digital Terrain Model (DTM).

The digital orthophoto is a processing final product suitable for cartographic use, of which the main characteristics are

- real and updated representation of the territory;
- geometric accuracy comparable with that of a technical map having a corresponding nominal scale;
- flexibility in processing and information system management.

Available digital orthophoto often derives from mosaicing many individually orthoprojected images.

The accuracy of an orthoimage determines its suitability for a certain scale map, which is its nominal scale.

If the orthoprojection is carried out on images acquired by remote sensing scanning systems, it is defined as an *orthodigital image*. In these cases, the sensor model (geometrical relationship between image and object spaces) is different from the equations of collinearity typical of frame images.

It is in this form that photogrammetry positively entered the world of remote sensing, based on scanning and radiometric processes, partially removing the barrier that saw the two techniques develop in parallel until now. Gradual improvement occurred in geometric resolution, making remote sensing close to photogrammetry. The fusion will be complete when scanning acquisitions are able to exhaustively reach analogical acquisitions' metric precisions and photographic process results.

Under the prompt of the availability of high-resolution images, suitable for cartographic updating at 1:10,000–1:5000 scales, more or less rigorous methods of image geometric reconstruction were developed, or borrowed from other fields.

Digital aerophotogrammetry has long since overtaken direct georeferencing, as soon as acquisition systems, made of a digital camera and auxiliary instruments for continuously measuring position and attitude, such as GPS/IMU integrated systems, appeared.

3.8 Oblique Photographs

Georeferenced oblique imagery extends the benefits of traditional straight-down photography providing a unique perspective view of a locality, allowing users:

- to see sides of a building, structure or feature, exposing blind spots, exits and entrances previously impossible to locate on straight-down photography;
- the ability to measure the height, length and area of features directly from photography;
- measurements can be made between real-world objects rather than their graphic representation in a 3D model;
- improve the identification of hard to see assets and facilities (e.g. lamp-posts, telegraph poles) which can be difficult to distinguish on traditional orthophotography;
- view GIS data in 3D by draping it on oblique imagery, extending the traditional and more familiar 2D view afforded by most GIS applications.

It is not intended that the oblique imagery replace existing orthophotography but rather compliment it.

The oblique photographs show buildings, infrastructure and land from all sides. This result in more detail, because there are multiple perspectives, with overlap resulting in as many as 10–30 images of the same location.

The most important applications are for planning and development, emergency response and property assessment. It also has applications in insurance, real estate, civil engineering and utilities.

Oblique photographs are a reference tool for occasions when certified measurements are not required but accurate information is needed to save time, resources and even lives – in an emergency situation, for example, when centimetre accuracy is not necessary.

3.9 Satellite Sensors for Photogrammetric Application

Satellite high resolution, thus with many limitations, is nowadays a possible alternative to aerial photogrammetry for the production of medium-scale (1:10,000–1:5000) Technical Cartography, a product hard to be generated and requiring continuous updating. This need requires the evaluation of production techniques taking into account costs/benefits analysis and possible solutions for exhaustive restitution.

Satellite high-resolution missions (IKONOS, QuickBird, OrbView 3, EROS, SPOT five satellites) nowadays perform geometric resolutions even better than 1 m that makes images suitable for Technical Map production at the scales previously stated.

The fact that these systems can carry out stereoscopic acquisition reduces the gap between them and classic photogrammetry although many doubts about the rigorous cartographic use of satellite data persist.

Unlike aerial photograms, satellite images do not follow the rigorous geometry of the central perspective, unique for the whole scene, and the problem of their correct orientation, especially of stereoscopic pairs, has not yet been completely solved.

The acquisition geometry characterizing them is typical of linear array, or push-broom, systems, such that each line of the scenes is individually acquired according to its own unique central perspective (see Chapter 6), generating geometrical distortions usually not considered in classic photogrammetry (see Chapter 8).

Once orientation problems are solved, the use of satellite images can be directed (like in classic photogrammetry) to:

- the production of orthophotos (geometrically corrected and georeferenced images);
- the realization of stereoscopic model correctly oriented for restitution processes;
- map planimetric updating, exploiting the orthophotos derived from high-resolution images.

Remote sensing advantages and disadvantages are discussed in Chapters 6 and 8.

3.9.1 Parametric Approach

The rigorous modelling of the geometric correspondences, or 'parametric approach', needs satellite position and attitude parameters (external orientation) and sensor optic–geometric parameters (internal orientation) to be known. As a consequence of the acquisition dynamics, they must be known for each scanning line, or block of lines to be considered geometrically homogeneous; GPS/IMU combined measures on board the platform can produce these data. The use of control points to reinforce the solution is desirable for correct estimation of external orientation parameters. Other distortions depend on the terrestrial curvature and on the adopted cartographic projection.

This approach (Toutin, 2002) is the most reliable and ensures the best accuracies. But if necessary data or indication about their correct meaning and accuracy are not provided, as happens too often, an alternative non-parametric approach needs to be adopted.

3.9.2 Non-parametric Approach

In a non-parametric approach, the correspondence between image space and object space is analytically modelled using methods chosen in relation to the characteristics of the areas to be processed (Box 3.6).

Box 3.6 The non-parametric approach requires the acquisition of several Ground Control Points (GCPs)

Method	Specification
Polynomial 3D	Small surfaces, several GCPs, corrections around the GCPs, low robustness
Rational Polynomial 3D (RFM)	Valid modelling, more similar to the parametric models, some problems of numerical instability
Thin Plate Spline (TPS)	Valid over lane areas, otherwise several GCPs are requested

Due to the high number of parameters to be estimated, in this case it is necessary to adopt many GCPs. These models are independent of the kind of platform and sensor considered. The most used model is the polynomial ratios one (RFM, Rational Functional Model), and in relation to this images distributors often provide RPC (Rational Polynomial Coefficients), i.e. polynomial coefficients taking part in the model. In this case, the operator is not required to collimate the GCP.

Non-parametric orthoprojection procedures allow the generation of orthophotos, from high-resolution images, suitable for cartographic updating up to 1:10,000 scale, with the chance to reach 1:5000 scale if GCP and DEM are very good quality.

3.10 Summary

The origins of photogrammetry go back to 1859, when photography was already well developed and studies about colour photography were underway. The purpose of this new discipline was to exhaustively describe the dimensions of an object and thus define its coordinates, which was done based on the intersection of a pair of photographic images.

Traditional photogrammetry uses images recorded on photographic paper. The reconstruction of geometric correspondences between object and image was firstly analogical and later analytical. A photogram acquisition geometry is determined by fixed parameters typical of the photo camera (which define the internal orientation of the photogram) and by parameters describing the relationship between the projection centre of the image and the reference system (external orientation). The latter, if not directly determined by GPS (direct georeferencing) which is quite a recent technique, are indirectly estimated by control points collected on the ground and identified on the photograms (indirect georeferencing).

The photogram is the output of the acquisition phase, realized by means of a photo camera provided with appropriate lenses. The photogrammetric camera, in the case of aerial acquisition, is fixed on an aircraft and runs parallel strips as defined in the flight plan in order to optimize the performance. Then in the restitution phase, the 2D photograms obtained are observed pair by pair through a binocular vision following specific rules, in order to have a stereoscopic view which perceives the third dimension. Different types of photogram orientation exist, each one requiring a different number of control points known and based on different equations.

Digital photogrammetry operates by using digital images (directly acquired by a digital camera or obtained by scanning traditional photograms), which presents many advantages in terms of resolution, time of availability of the images and costs compared with analogical images. In photogrammetry the most used format of digital images are TIFF and JPEG. Photograms' orientation and restitution are operated directly by software accessible by a large amount of averagely specialized users.

An aerial photogram is a central projection of the portrayed portion of land and needs to be transformed into an orthographic projection by correction procedures in order to be used for cartographic production.

High geometric resolution performed by the latest satellites and the possibility to acquire stereoscopic images nowadays offer an alternative to aerial photogrammetry for medium-scale cartographic production, although correction procedures are always required to correct geometric distortions.

Further Reading

Cliff G., (Ed.), 1996, Digital Photogrammetry: An Addendum to the Manual of Photogrammetry. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland.

Ebner, H., Fritsch D., Heipke C. (Eds.), 1991, Digital Photogrammetric Systems. Wichmann Verlag, Karlsruhe.

Kraus K., 1993, Photogrammetry. Vol. 1. Fundamentals and Standard Processes. Dummler, Bomm, p. 397.

Li Z., Chen J., Baltsavias E. (Eds.) 2008, Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences, 2008 ISPRS Congress Book. CRC Press, Taylor & Francis Group, Boca Raton, London, New York, Leiden, p. 527.

Mikhail E.M., Bethel J.S., McGlone J.C., 2001, Introduction to Modern Photogrammetry. John Wiley & Sons Inc., New York.

Bibliography

Hinz A., Dorstel C., Heier H, 2000, DMC 20001 System concept and data processing work-flow. GIM International, August: 45–47.

Heipke C., 2002, Digital Georeferencing in Photogrammetry – The future of Airborne Surveying, Atti Seminario Internazionale La Fotogrammetria nell'era inerziale, 13–14 June. SIFET, Pavia, Italy.

Jacobsen K., 2007, Comparison of large size digital airborne frame cameras with analogue film cameras. Proceedings of the 27th EARSeL Symposium, GeoInformation in Europe, M.A. Gomarasca (Ed.). Millpress, The Netherlands, pp. 89–96, ISBN 9789059660618.

Kalman R.E., 1960, A new approach to linear filtering and prediction problems. Transactions of the ASME – Journal of Basic Engineering, 82: 35–45.

Lehmann F., Hoffmann A., Renouard L., van der Vegt J.W., 2000, The high resolution stereo camera-airborne (HRSC-A). GIM International, July: 12–17.

Loedeman J.H., 2000, Three-line linear versus multi-head array. GIM International, May: 68–71.

McCurdy P.G., et al. (Eds.), 1944, Manual of Photogrammetry. Preliminary edition. Pitman Publishing Corp., New York.

Paine D.P., 1981, Aerial Photography and Image interpretation for Resource Management. John Wiley & Sons, New York, p. 571.

Philipson W.R., 1997, Manual of photographic interpretation. American Society for Photogrammetry and Remote Sensing, 2nd Edition: 3–20.

Slama C.C., Theurer C., Henriksen S.W. (Eds.), 1980, Manual of Photogrammetry, 4th ed. American Society for Photogrammetry and Remote Sensing, Falls Church, Virginia.

Toutin Th., 2001, Elevation modelling from satellite visible and infrared (VIR) data: a review. International Journal of Remote Sensing, 22(6): 1097–1225.

Toutin Th., 2002, Three-dimensional topographic mapping with ASTER stereo data in rugged topography. IEEE Transactions on Geoscience and Remote Sensing, 40(10): 2241–2247.

Wolf P.R., 1988, Elements of Photogrammetry, Cap. 3: Principles of Photography. McGraw-Hill International Editions, McGraw-Hill Book Company, International Editions, New York, pp. 41–60.

Chapter 4
Elements of Remote Sensing

The main goal of *Remote Sensing* (RS) is the production of maps, mainly thematic but also topographic. Its fundamental characteristics as a source of information for georeferencing are

- *synoptic vision* of ground surface conditions otherwise not obtainable with traditional techniques;
- *repeating cycle*: periodic observations enabling temporal comparisons and *updating* of the collected data;
- *multispectral* acquisitions.

Information acquisition with techniques of remote sensing is developed in three phases (Fig. 4.1):

- *collection of data* from the ground, aerial and/or satellite acquisition stations;
- *processing* of the collected data;
- *data interpretation*, followed by the restitution, hardcopy or digital, of thematic cartography.

From the instruments used for collecting the information, it is possible to derive

- *measures*: using radiometers, spectrophotometers, scatterometers;
- *images*: obtained by traditional and digital photographic cameras, or photo cameras, scanners, thermal cameras and radar.

The instruments are divided into:

- *passive sensors*: recording the intensity of the reflected electromagnetic energy coming from the Sun or emitted by the Earth: photo cameras, scanners, thermal cameras and video cameras;
- *active sensors*: the acquisition systems emit radiation themselves, collecting the back signal (radar); the radiation is backscattered to the sensor with intensity depending on the structural characteristics of the examined surface and on the wavelength (λ) of the incident energy.

M.A. Gomarasca, *Basics of Geomatics*, DOI 10.1007/978-1-4020-9014-1_4,

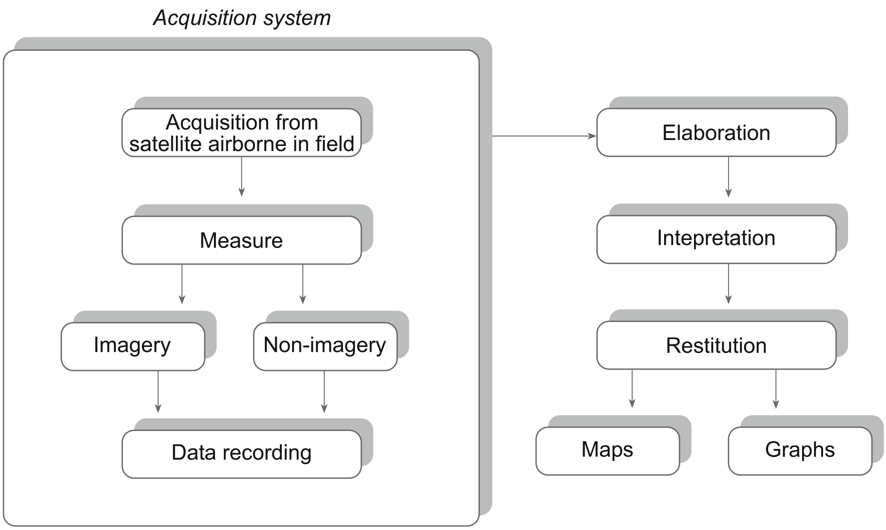


Fig. 4.1 Remote sensing process phases and restitution of the results

It is also common to distinguish between

- *optical*: spectral range in the interval 0.3–15 μm, typical of passive remote sensing, identified by the sensors:

– *panchromatic*: one band including the visible range and in some cases part of the near infrared;
– *multispectral*: 2–9 spectral bands;
– *super-spectral*: 10–16 spectral bands;
– *hyperspectral*: more than 16 spectral bands;

The increase of the number of bands in general improves the bandwidth (bandwidth is more common) and the spectral interval.

- *radar*: microwaves ranging from 1 mm to 1 m, typical active remote sensing tool, that can operate, with single or multi-polarization and with single or multiple incidence angle, in:

– single frequency;
– multi-frequency.

Technical problems of acquisition and representation are currently being operatively solved, while problems still exist in understanding the characteristics of these techniques by decision makers and administrators at national, regional, provincial and city level.

4.1 Milestones in the History of Remote Sensing

Like aerial photogrammetry, remote sensing derives from photography (Chapter 3), optics and spectrometry; its history is deeply entwined with the domains of electromagnetic spectrum and aeronautics.

Galileo Galilei (1564–1643) and Isaac Newton (1642–1727) founded optics and stereoscopic science. In ancient times, Aristotle in 300 BCE conducted experiments with the camera obscura and Seneca (4 BCE–64 CE) considered optics and refraction. More recently Leonardo da Vinci (1452–1519) through experiments discovered the possibility of splitting white light into its colour components. He recognized five out of the seven traditional iris colours, but kept his experiments secret avoiding discrediting the universally acknowledged purity of white light, and probable persecution. The phenomenon was so mysterious that it was defined as the *light spectrum* by Leonardo.

In 1666 Isaac Newton carried out experiments and described in a scientific way the known phenomenon of a white light beam passing through a triangular prism that makes it split into the seven iris colours: violet, blue, cyan, green, yellow, orange and red (see Plate 4.1); using a second prism Newton demonstrated that the colours can be recombined into white light.

Until the end of the XVIII century, it was commonly thought that sunlight is totally perceived by the human eye and therefore the emission was limited to what is nowadays defined as the visible spectrum. But the Anglo-German astronomer and musician William Herschel, at that time popular for discovering the existence of the planet Uranus, in 1800 demonstrated that sunlight passing through different coloured filters produces different temperature levels. He arranged an experiment using Newton's prism to split white light into the iris colours and measured the temperature of each colour.

Box 4.1 Reference year, scientist, and scientific discovery in the fields of the electromagnetic spectrum and remote sensing theory

Year	Scientist	Discovery
1609	Galileo Galilei	Optics
1666	Isaac Newton	Spectrometry
1669	Erasmo Bartolini	Light polarization
1676	Olaus Roemer	Calculation of the speed of light
1690	Christian Huygens	Theory of the refraction based on the wave nature of the radiation
1729	James Bradley	Light aberration
1753	Thomas Melvill	Emission of different colours of different chemicals on a flame
1758	John Dolland	Achromatic lens
1783	William H. Wollaston	Use of a hole for the spectral study of the Sun

1800	William Hershel	Infrared
1801	J. W. Ritter	Ultraviolet spectrum
1801	Thomas Joung	Theory of the perception of the colour
1810	Etienne L. Malus	Theory of the double refraction
1819	A. J. Fresnel	Diffraction paths in the hypothesis of the light as a wave phenomenon explained in his theory
1822	Thomas J. Seebeck	Thermoelectric effect
1828	William Nicol	Particular prism able to produce polarized light
1829	Leopoldo Nobili	First thermocouple base on the experience of Seebeck
1839	Louis J. Daguerre	Photography
1845	Michael Faraday J. B. L. Foucault	Relationship between magnetism and light Extension of the measures of the solar spectrum till 1.5 μm
1857	J. B. L. Foucault	Demonstration that the light velocity changes in different media
1858	Gustav R. Kirchhoff	A good absorber is also a good emitter of energy Development of the first spectroscope
1868	Alexandre E. Becquerel	Effects of the near infrared in the field of the photography and of the phosphorescence
1872	John Tyndall	Measure of the radiant power transferred from a heated body to an other
1881	Samuel P. Langley	Extension of the measure of the solar spectrum till 18 μm
1884	T. Stefan, L. Boltzmann	Stefan–Boltzmann's law
1886	Abraham Michelson	Precise measure of the velocity of the light. He invented his interferometer with the purpose of discovering the effect of the Earth's motion on the observed velocity
1900	Max Planck	Revolutionary idea that the energy emitted by a resonator could only take on discrete values or quanta. The energy for a resonator of frequency v is hv where h is a universal constant, now called Planck's constant
1904	Hulsmeyer	Demonstration of the use of the microwaves for object detection
1936	Robert Watson-Watt	First pulse radar
1946	Several authors	First microwave images of the Earth's surface
1954	Several authors	Side Looking Aerial Radar (SLAR)
1964	Several authors	Synthetic Aperture Radar (SAR) based on the Doppler effect
1972	NASA-ERTS-1 Landsat	First satellite for Earth Observation

Herschel described the increase in temperature as going from violet to red and observed, by chance, that the temperature was even higher beyond the red colour, where nothing was supposed to be. He deduced the existence of *invisible heating rays* that were until then unknown. One year later J. W. Ritter discovered the

existence of ultraviolet light: the blackening of few silver salts put beyond the blue-violet line allowed him to hypothesize the presence of energy, in the specific case of photochemical energy, further than the visible in the opposite side to the red. The year 1873 is one of the most important in the history of the light spectrum as James Clark Maxwell presented his mathematical theory about electromagnetic radiation. The succession of events leading to the modern conception of the electromagnetic spectrum is very wide. The list of scientists, together with the year of reference and the discovery to which their name is linked, is reported in Box 4.1.

In 1887 Heinrich R. Hertz produced very long infrared waves with an electrical device; so the similarity between electromagnetic waves produced by a magnetic field and by an electrical field became clear. At that time Hertz's studies contributed to diminishing the radiation theory based on particles or photons, affirming the concept of electromagnetic waves.

In 1900 Plank's radiation law was still merging the radiation *corpuscular theory* with the *undulation theory*.

As a matter of fact, the principles of remote sensing are based on both the undulation theory by Huygens and the corpuscular theory by Newton.

4.2 Electromagnetic Spectrum

Electromagnetic (EM) spectrum (see Plate 4.2) is mono-dimensional continuous, consisting of a set of radiations ordered according to wavelength, frequency or photonic energy; it includes waves of any wavelength, ranging from fractions of Angstrom (one tenth of a millimicron $Å = 10^{-10}$ m) to many kilometres. As a matter of fact, there is not one individual source or detection system suitable for the whole electromagnetic spectrum, which is thus divided into spectral regions, according to the instruments for generating, isolating and detect radiation. EM energy can be compared with the wind: it is detectable only by the effects it produces.

For practical use, the electromagnetic spectrum is divided into regions, according to wavelength (Fig. 4.2 and Table 4.1): *visible light*, so defined as its range is within the interval of the human eye's spectral sensibility; cosmic rays, gamma (γ), X-rays, ultraviolet, infrared, visible, microwaves, radio waves – these are the most well-known energy types existing in nature. This energy has some constant characteristics and radiates according to the undulation theory which describes EM energy as waves propagating in a vacuum in harmonic motion, i.e. by regular intervals in time, and at light speed (Fig. 4.3).

Electromagnetic radiation consists of an electrical field (E) that varies in magnitude along a direction perpendicular to the direction of propagation. The magnetic field (H), oriented at right angles to the electrical field, is propagated in phase with the electrical field. It is described by

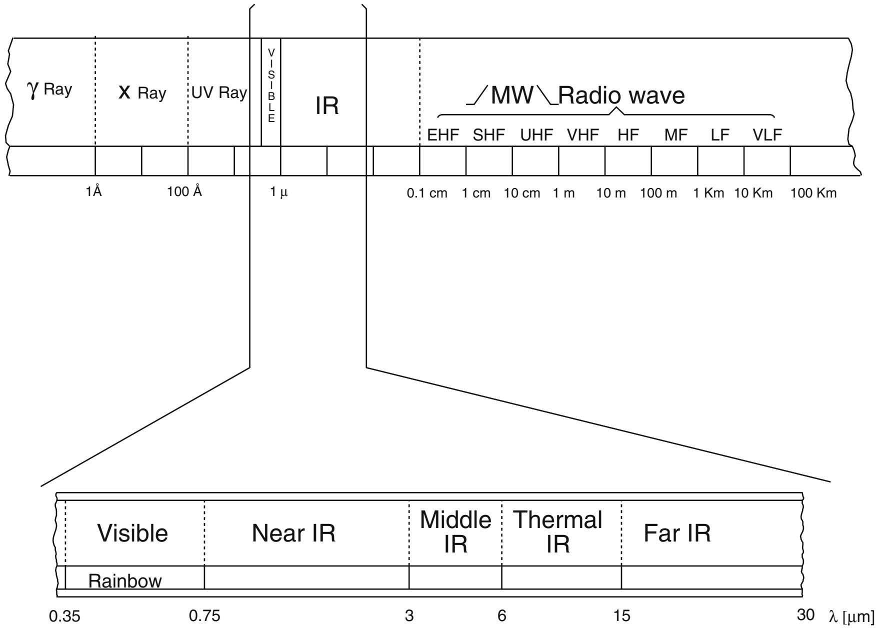


Fig. 4.2 The regions of the Electromagnetic Spectrum: wavelength (λ) is inversely proportional to frequency (ν). In the visible, the rainbow colours starting from the shorter λ are *violet*, *blue*, *cyan*, *green* and *yellow*, *orange* and *red*

- *wavelength* (λ): measured in microns (μm), is the distance from one peak, or crest, of a wave of energy to the next corresponding peak or crest;
- *frequency* (ν): is the measure of the number of occurrences of repeating EM waves per unit of time. This is described as the number of waves made per second, measured in cycles per second or hertz (Hz).

Table 4.1 Spectral regions of interest in remote sensing

Radiation	Wavelength (λ)	
Ultraviolet (UV)	0.01–0.38 μm (10–380 nm)	
Visible (V)	0.38–0.75 μm (micron)	
Violet		0.380–0.455 μm
Blue		0.455–0.500 μm
Green		0.500–0.580 μm
Yellow		0.580–0.595 μm
Orange		0.595–0.620 μm
Red		0.620–0.750 μm
Infrared	0.75–1.0 mm	
Near (VIR)		0.75–3.0 μm
Medium (MIR)		3.0–6.0 μm
Far or thermal (TIR)		6.0 μm–1.0 mm
Microwave (MW)	0.1–100 cm	

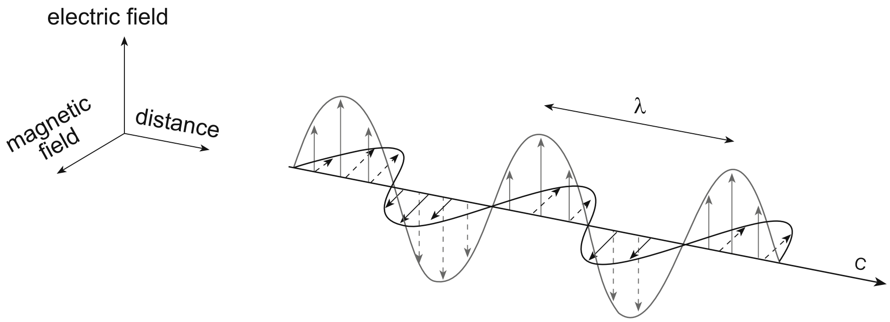


Fig. 4.3 Propagation of electromagnetic energy occurs at the light velocity c following a defined model of the undulatory (wavelike) theory: the wavelength λ is defined as the average distance from one wave crest to the next, and it is proportional to the frequency ν

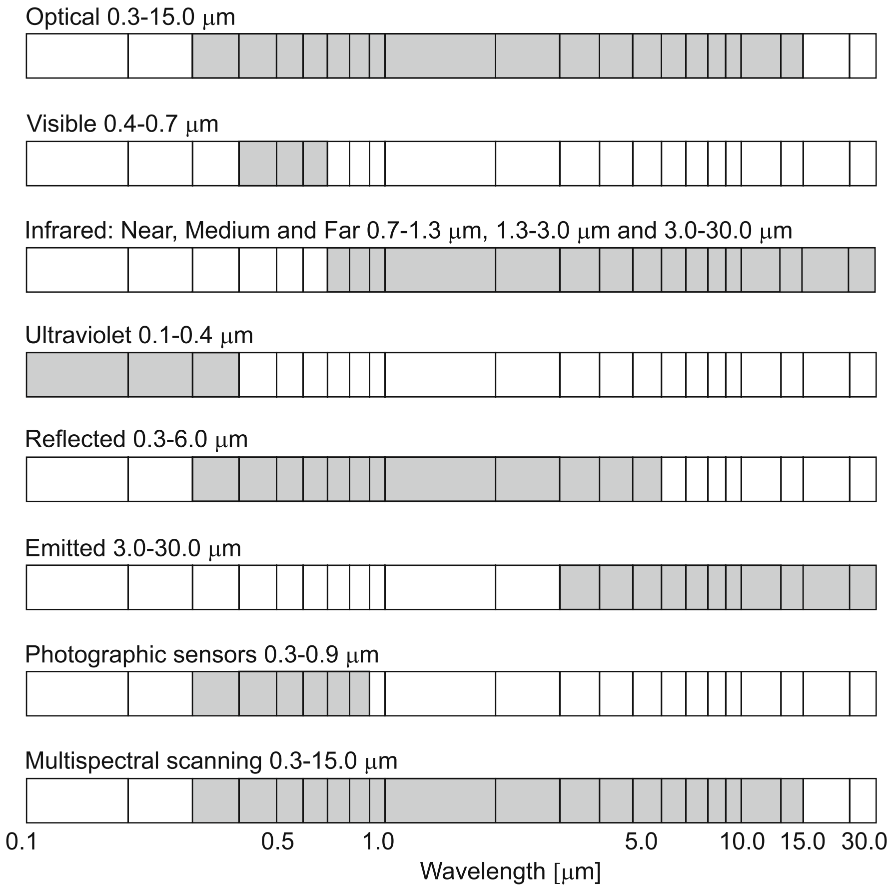


Fig. 4.4 Spectral intervals of the optical remote sensing and sensitivity of the photographic and digital sensors

The electromagnetic spectrum is considered as a continuum of electric, magnetic and visible radiation, ordered according to their frequency, wavelength or wave number.

In remote sensing, there are two typologies of instruments, passive and active, based on the source of incident radiation on the Earth's surface and on wavelength intervals.

In passive remote sensing, sensors operate in wavelength intervals from ultraviolet to thermal infrared; in active systems such as radar, they operate in microwave intervals (Fig. 4.4).

Multispectral acquisitions in the far infrared range 10.4–12.5 μm are defined as thermal. *Thermal* means that in this band the radiance detected by the sensors is energy emitted by the surfaces, according to Planck's law, as a function of the surface's temperature and on the emissivity coefficient ε.

4.3 Optical Passive Remote Sensing

4.3.1 Sources of Electromagnetic Waves: The Sun and the Earth

The Sun and the Earth are complex physical bodies. The studies about radiation phenomena which characterize them are fundamental for the application of remote sensing techniques.

The Sun can be represented as a sphere of incandescent gas whose apparent surface (photosphere) has a diameter of about 1.39×10^6km. The electromagnetic radiation emitted by the Sun ranges from cosmic rays to radio waves; the visible spectral range is about 50% of the energy emitted by the Sun and reaches the Earth's surface in about 8 min and 30 s, illuminating and heating it. This energy coming from the photosphere is almost constant (variations lower than 1% are recorded) corresponding at 2 small calories per minute per cm^2 of Earth surface (1000 small calories correspond to 1 kilocalorie or big calorie) perpendicularly exposed to the Sun's rays (solar constant).

The total radiation (integrated over the whole electromagnetic spectrum), incident on a surface normal to the Earth–Sun axis in the vacuum space at an average Earth–Sun distance, has a value of about 1377 W/m^2 (defined as solar constant) with a standard deviation of 8 W/m^2.

The flux of electromagnetic radiation emitted by the Sun is complicated by strong temperature variations occurring between the centre and the surface and by the relative opacity of some regions of the atmosphere at different wavelengths. Measuring the Sun's energy at a given wavelength, the apparent blackbody temperature of the Sun can be obtained inverting the Planck's law; changing the wavelength, the temperature fluctuates from a minimum of 4500 K to a maximum of 9000 K, with an average value of 6000 K.

If the photosphere temperature is 6000 K and considering the behaviour of the Sun similar to the surface of a blackbody, according to Wien's law, it can be seen

that this energy is mostly irradiated (48%) within the visible range, from 0.4 to 0.7 μm wavelength (see Plate 4.3).

The Earth, with an average surface temperature of about 288 K (15°C), mostly radiates in the thermal infrared band, ranging from 8 to 14 μm with the potential peak around 9.6 μm; the side exposed to the Sun reflects part of the incident solar light.

In the spectral range from 2.5 to 5 μm, the average energy emitted by the Earth's surface can be compared with the average energy reflected by the Earth's surface but coming from the Sun. Detection and analysis of this interval therefore requires particular attention.

The typical curve of exitance of the Earth's surface is lower with respect to the Sun curve, due to the medium temperature value of about 300 K.

Electromagnetic radiation coming from Sun undergoes attenuation due to the following main causes:

- Earth–Sun distance;
- absorption and scattering processes characterizing the atmosphere.

Earth–Sun distance is about 149.6×10^6km; as the intensity of the incident radiation decreases, according to the inverse quadratic distance between the two bodies, the Sun's energy has a reduction of 2.16×10^{-5}, i.e. about 1/200,000 of the energy emitted by Sun reaches the external limit of the atmosphere (Table 4.2).

Table 4.2 Solar spectral irradiance in percent at the external limit of the terrestrial atmosphere

Wavelength (λ) (μm)			Solar spectral irradiance (%)
Ultraviolet	UV	$\lambda < 0.38$	9.5
Visible	VIS	$0.38 < \lambda < 0.76$	45.7
Near infrared	NIR (VIR)	$0.76 < \lambda < 2.2$	38.3
Medium infrared	MIR	$\lambda > 2.2$	6.5

The incident solar energy on the Earth's surface changes in function of the Sun's angle; at 40–45° of latitude it is maximum in June, while it is minimum in December (Plate 4.4).

4.3.2 Physical Principles

Max Planck, a German physicist honoured with the Nobel Prize in 1919, formulated a theory in 1900 that introduced into atomic physics the postulation for which a light source does not emit energy in a continuous but in a discontinuous way, contrasting with the certainties of the period. He hypothesized the existence of granules of energy, called *quanta*, as quantity or dose. This theory was a true revolution that led to the development of *quantum mechanics*, built on the discoveries of Heisenberg, de Broglie and Schrodinger, scientists from the German school who received

the Nobel Prize for Physics for their studies on the mechanics of matrices and on the undulation theory of light. Before Planck, the Universe was considered a perfect mechanism with defined actions and reactions; after these studies, it was represented as an entity unpredictable in its movements and in the results.

This theory makes space and time indivisible and admits the possible coexistence in different physical states of an entity like light that is at once wave and corpuscle.

People interested in some aspect of the physical sciences regularly use quantum physics and mechanics. Without it semiconductors, electronic computers, the discovery of DNA or much progress in diagnostics and therapy in medicine would not have been made.

Remote sensing is based on electromagnetic energy properties. The fundamental laws of radiation defining its quantitative aspects are

- *Kirchhoff's radiation law:* regulates the relationship among the coefficients of reflection, transmission, absorption and emission;
- *Planck's radiation law:* defines the behaviour of the energy emitted by a surface as a function of wavelength and temperature;
- *Stefan–Boltzmann's radiation law:* furnishes the total quantity of energy emitted by a surface calculated on the whole electromagnetic spectrum, for any temperature;
- *Wien's displacement law:* points out the wavelength value in correspondence with the maximum electromagnetic emission at a defined temperature.

Since the physical laws at the base of remote sensing are valid for the ideal surfaces of a blackbody, it is opportune to describe the general characteristics of such a physical object.

A blackbody is intended as an ideal surface able to absorb all the incident radiation upon it and re-radiates energy according to Planck's laws. The properties of the emitted radiation are independent of the properties of the body, thus its spectrum is typically used for studying the emission of a continuous spectrum.

In 1860 the German physicist Kirchhoff demonstrated that a body acts as a perfect blackbody if

- the sides of the body are maintained at a constant absolute temperature (temperature of blackbody);
- a very small hole in comparison to the dimensions of the body is made in the body itself.

No electromagnetic radiation passes through it and none is reflected. Because no light is reflected or transmitted, the object appears black when it is cold.

If the blackbody is hot, these properties make it an ideal source of thermal radiation. If a perfect blackbody at a certain temperature is surrounded by other objects in thermal equilibrium at the same temperature, it will on average emit exactly as much as it absorbs, at every wavelength.

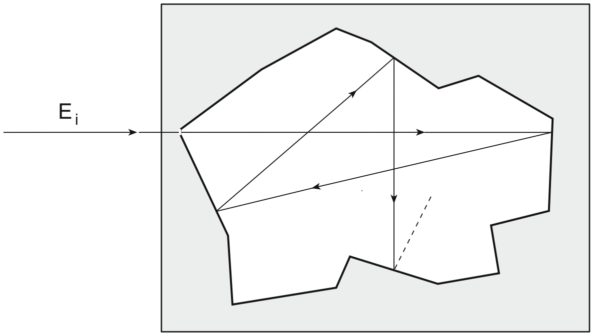


Fig. 4.5 Surface of a blackbody represented by a small hole in a cavity where the light passes through without any exit of energy and in which it is completely absorbed. The blackbody, a hypothetical source of energy, is an ideal body which would be a perfect absorber and a perfect radiator, absorbing all the incident radiation, reflecting none and emitting radiation at all wavelengths

This radiation is named *blackbody radiation*, or cavity radiation, and it depends on the blackbody temperature and not on the wavelength (Fig. 4.5).

Further properties of a blackbody are defined in the enunciations of the four fundamental laws of remote sensing.

4.3.2.1 Kirchhoff's Radiation Law

In his studies Kirchhoff started from the assumption that every radiation incident on a real surface reacts following three phenomena: *reflection*, *absorption* and *transmission*.

From the principle of the conservation of energy (Fig. 4.6), it is deduced that

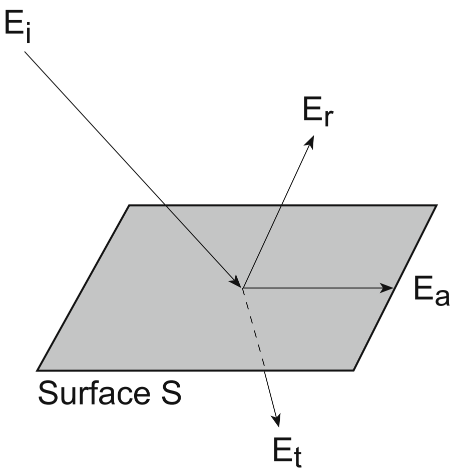


Fig. 4.6 Kirchhoff's law

$$E_i = E_r + E_a + E_t \tag{4.1}$$

where

E_i : incident radiation
E_r : reflected radiation
E_a : absorbed radiation
E_t : transmitted radiation.

Factors in Eq. (4.1) can be associated with coefficients that quantify the phenomena of absorption, reflection and transmission verified during the radiant energy–matter interaction. The incident energy (E_i) can be reflected (E_r), absorbed (E_a) or transmitted (E_t) in function of the following parameters:

Coefficient of reflection	$\rho = E_r / E_i$	$0 \geq \rho \geq 1$
Coefficient of absorption	$\alpha = E_a / E_i$	$0 \geq \alpha \geq 1$
Coefficient of transmission	$\tau = E_t / E_i$	$0 \geq \tau \geq 1$

where ρ, α and τ are the adimensional coefficients of reflection (or reflectivity), absorption (or absorptivity) and transmission (or transmissivity); ρ or reflectivity: ratio of the reflected energy by a surface to the incident energy; α or absorptivity: ratio of the absorbed energy by a surface to the incident energy; τ or transmissivity: ratio of the transmitted energy by a surface to the incident energy.

Their sum gives a unitary result:

$$\rho + \alpha + \tau = 1 \tag{4.2}$$

This relationship describes the most general case; transparent bodies only transmit ($\tau = 1$), or opaque bodies do not transmit at all ($\rho + \alpha = 1$). Bodies having a constant behaviour in all the wavelengths do not exist in nature. For this reason, the reference to a definite wavelength or *spectral range* (α_λ, ρ_λ, τ_λ,) is always needed.

In remote sensing, only the radiation reflected and emitted by objects is considered. If a body generates an isotropic reflection, i.e. uniform in all directions, the surface is called *Lambertian* (Fig. 4.7).

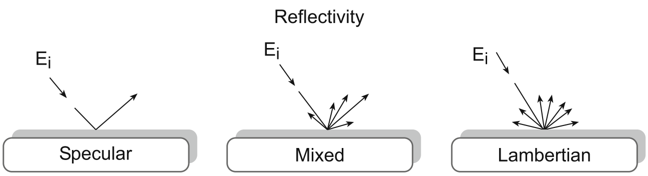


Fig. 4.7 Different situations of incident energy reflection

Before enunciating Kirchhoff's law, it is necessary to introduce the emissivity ε, expressing the efficiency with which a body irradiates and is defined by the following ratio:

$$\varepsilon = \text{emitted energy by the real body at } \mathrm{T_0}/\text{emitted energy by the blackbody at } \mathrm{T_0}$$

Since the blackbody absorbs all the energy, and at the same time emits radiation with higher efficiency, its value of emissivity is equal to 1.

Starting from these relationships, Kirchhoff was able to demonstrate that

$$\alpha = \varepsilon$$

If the coefficient of absorption α is equal to the coefficient of emission ε; it means that a good radiator is also a good absorber. Therefore in the equation relating the parameters of absorption, transmission and reflection, the emissivity ε can substitute the absorption:

$$\rho + \varepsilon + \tau = 1 \tag{4.3}$$

insofar as τ is equal and for opaque surfaces; the more the surfaces reflect, the less they emit electromagnetic radiation and vice versa.

4.3.2.2 Planck's Radiation Law

Any external surface of a body, if its temperature is higher than zero absolute expressed in degrees Kelvin (–273.14°C), emits electromagnetic radiation in relation to the body temperature and to the physical–chemical–geometric characteristics of its surface, while it reflects, absorbs or transmits the electromagnetic radiation coming from an external source. The two main sources of electromagnetic energy are the Sun and the Earth.

Electromagnetic radiation is defined by wavelength (λ) and by radiation frequency (ν), related by a relation of inverse proportionality:

$$\mathrm{c} = \lambda\nu \qquad [\mathrm{m \cdot s^{-1}}] \tag{4.4}$$

where c:light speed in a vacuum; λ : wavelength; ν: radiation frequency.

The speed propagation of the light, or of the electromagnetic radiation, can be considered constant in the vacuum and equal to about 300,000 km/s^{-1}.

The general law of electromagnetic emission was enunciated by Planck in December 1900 in synthetic form:

$$\mathrm{e} = \mathrm{h}\nu \qquad (\mathrm{W}) \tag{4.5}$$

where e: quantum of energy of the radiation; h: Planck quantum (a constant); ν : radiation frequency.

The equation points out how the quantity of energy provided by the radiation is directly proportional to the frequency and inversely proportional to the wavelength.

One of the possible mathematical formulations of Planck's radiation law is expressed by the equation:

$$M_\lambda = \frac{2\pi hc^2\, h\lambda^{-5}}{e^{\frac{ch}{\lambda kT}} - 1} \qquad [Wcm^{-2}\mu m^{-1}] \tag{4.6}$$

where
M_λ: spectral radiant exitance per surface unit and wavelength unit [$W \cdot cm^2 \cdot \mu m$]
λ : radiation wavelength [μm]
π 3.1415
c : 2.99×10^{10} ($cm \cdot s^{-1}$) light speed in a vacuum
h : 6.62×10^{-34} ($W \cdot s^2$) Planck constant
k : 1.38×10^{-23} ($W \cdot s \cdot K^{-1}$) Boltzmann constant
e : 2.718281828459045 quantum of energy of the radiation
T : absolute radiant temperature (K) 0 K = −273.14 °C

The energy emitted by a blackbody, at a definite wavelength, is a function of the temperature of the body. A blackbody has unitary emissivity ($\varepsilon = 1$). The density of radiant flux $d\Phi/dA$ is called *exitance* (M_λ) if considering the exiting flux from a surface and *irradiance* (E_λ) when incident on a surface (see Box 4.2).

4.3.2.3 Stefan–Boltzmann's Radiation Law

The Stefan–Boltzmann law gives the total energy being emitted at all wavelengths by the blackbody, which is the area under the Planck Law curve. Thus, the Stefan–Boltzmann law explains the growth in the height of the curve as the temperature increases.

$$M = \int_{\lambda_1=0}^{\lambda_2=\infty} M_\lambda d\lambda = \sigma T^4 \qquad [Wcm^{-2}] \tag{4.7}$$

where M : radiant energy per surface unit; M_λ : radiant energy per surface unit and wavelength unit; σ : 5.67×10^{-12} ($W \cdot cm^{-2} \cdot K^{-4}$) Stefan–Boltzmann constant; T : absolute temperature of blackbody in degrees Kelvin (K).

According to this equation, the radiation emitted by a body surface is a function of the fourth power of its absolute temperature. This law directly relates the total exitance of a surface and the value of its absolute temperature, independent of the wavelength.

Stefan–Boltzmann's law is valid only for the surfaces of the blackbody, with $\varepsilon = 1$. In reality, bodies have emissivity from 0 to 1 and are a function of λ, therefore the law, if applied to non-blackbodies, assumes the form:

$$M_\lambda = \varepsilon(\lambda) \cdot \sigma T^4 \tag{4.8}$$

4.3.2.4 Wien's Displacement Law

Planck's law is a function with a maximum value for any temperature (Fig. 4.8) but, before it was defined, Wien enunciated the law, in 1893, that relates the wavelength λ, corresponding to the maximum of energy emitted by a blackbody surface, with its temperature value T:

$$\lambda_{\max} = \mathrm{w} \cdot \mathrm{T}^{-1} \quad (\mu\mathrm{m}) \tag{4.9}$$

where w : specific constant for each body according to its emissivity characteristics; for a blackbody w = 2897.8; T : absolute temperature of blackbody in degrees Kelvin (K).

T is the temperature that a blackbody surface would have if it emitted like the surface of reference.

Wien's displacement law enunciates that there is an inverse relationship between the wavelength of the peak of the emission of a blackbody and its temperature, in accordance with Planck's law. For instance, a red flame, λ: 0.72 μm, has smaller caloric input in comparison to the blue-violet flame, λ: 0.40 μm. This law is useful in practical terms because the optimal spectral band of investigation for a surface at a definite temperature can be selected.

The Sun, for example, having a blackbody apparent temperature of about 6000 K, has its maximum emission at 0.483 μm corresponding to blue-green light; the objects at 25 °C (298 K) emit around 10 μm. The Earth's surface emits at 9.66 μm, in the thermal infrared.

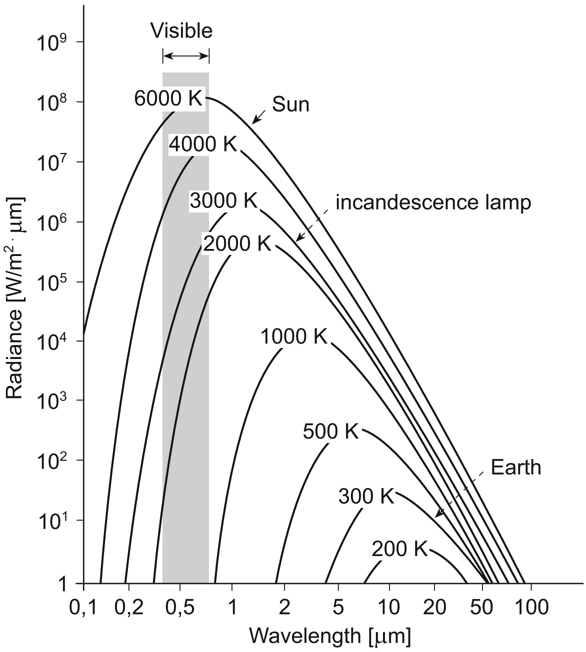


Fig. 4.8 Radiance curves of a blackbody at different temperatures (Plank's law). The wavelength moves from the maximum of emission towards shorter wavelengths as temperature increases (Wien's law). The area delimited by each curve is the total radiance emitted at that temperature (Stefan–Boltzmann's law). The Sun, T: 6000 K, has a peak of emission at 0.483 μm (0.4–0.7 μm visible range); the Earth, T: 300 K, has a peak of emission around 9.66 μm

4.3.2.5 Terrestrial Albedo

The *albedo* is defined as the ratio of the electromagnetic energy reflected or diffused by a surface to the total incident energy. The albedo is therefore equivalent to the *reflectivity* ρ of a body compared with the reflectivity of a surface positioned at the same distance from the Sun, perpendicular to it and diffusing the whole incident radiation. This relationship is usually expressed as a percentage.

Table 4.3 Albedo values for different surfaces

Surface	Albedo (%)
Coniferous	10–20
Broadleaf	15–25
Cultivated field	15–30
Light soil (sand)	35–40
Dark soil (organic soil)	5–10
Desert	25–30
Recent snow	75–95
Water (incidence angle < 45°)	5–10
Water (incidence angle > 45°)	5–99

Albedo of different objects varies as reported in Table 4.3. Terrestrial albedo is approximately 34%. The values are comparable if referring to the global radiation in nadir hours of a cloudless day. The presence of partial cloud cover considerably modifies the values of albedo compromising the normalization and the use of the data.

The quantity of energy available for ecosystems is higher in comparison with smaller values of albedo.

The albedo, in vegetation, is a function of the phenological phase and therefore of fractional cover and biomass density. High-density vegetation with low trees gives higher values of albedo than sparse woodland with high trees; diffusion and absorption increase in the canopy and in the lower layers. The values of albedo vary with the Sun's angle. The variations are larger with tilted angles of the solar rays from 0° to 30°. A healthy biomass observed at different hours may vary the albedo value from 30% with a Sun–horizon angle of 10°, to 15% with a Sun–horizon angle of 60°.

4.3.3 Visible Radiation and Colour

Colour is the visual perceptual property in perceiving the external world physically originated by coloured lights and/or coloured pigments. It interacts in the eye with the spectral sensitivities of the light receptors. In the sensitive organs (receptors) of the visual apparatus, the biochemical stimulus is generated producing impulses then elaborated by the brain. Biochemical stimulus is induced by a solicitation of the electromagnetic energy perceived in the range of the visible wavelength from 0.4 to 0.7 μm.

The tristimulus is the model of reference for the description of colorimetric perception based on the following assumptions:

- the retina of the eye, behaving like a photographic film, contains cellular systems, the *rods*, that act as photo-detectors, sensitive to the luminous stimulus (also low-intensity ones) able to transmit achromatic images, and three types of *cones* sensitive to red, green and blue wavelengths (Fig. 8.33);
- chromatic perception and vision are possible thanks to the combination of these signals transmitted to the brain by the optic nerve.

For the definition of a standardized colorimetric system, it has been necessary to understand the relationship among the reactivity of the three photo-receptors of the eye during the perception of a colour in order to draw the functions of chromatic, or spectral, sensibility.

The experiments conducted by Stiles–Burch and Speranskaya have been resumed in three curves, called functions of chromatic correspondence (tristimulus).

The tristimulus system is based on visually matching a colour under standardized conditions against the three primary colours red, green and blue; the three results are expressed as X, Y, and Z, respectively, and are called tristimulus values (see Plate 4.5). Such functions represent the key to realize the transfer between the external physical reality of the colour and the internal perceptive reality of the human eye (Fig. 4.9).

Tristimulus diagrams express the relationship among the spectral characteristics of the three monochromatic primary colours selected to define the reference system; the primary colours have been fixed to the wavelength of 645.2 nm, red, 526.3 nm, green and 444.4 nm, blue. If the references changed, the results of the syntheses

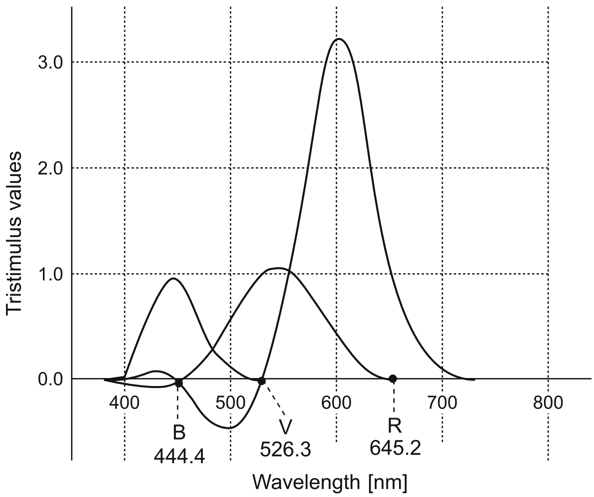


Fig. 4.9 The average tristimulus functions resulting from the Stiles–Burch and Speranskaya experiments

would be different; thus, the choice of the primary colours is important and based on the fact that each primary colour is not reproducible with the synthesis of the other two.

The first experimental results showed some limits, due to the functioning of the cones in the human retina, among which

- the impossibility of obtaining the white colour with tristimulus weights equally distributed in R, G, B;
- the presence of negative values of tristimulus in the spectral range 450–550 nm.

To overcome such incongruities, the CIE (*Commission Intemationale de l'Éclairage*) in 1931 fixed the experimental curves in a new primary colours system:

R : 700.0 nm
G : 546.1 nm
B : 435.8 nm.

The problem of the negative values was solved by defining linear algebraic transformations of values that produce imaginary colours called X, Y and Z. The advantage is the expression of natural and artificial colours in positive values. The peaks of the three functions correspond to 0.600, 0.555 and 0.450 μm wavelength (Fig. 4.10).

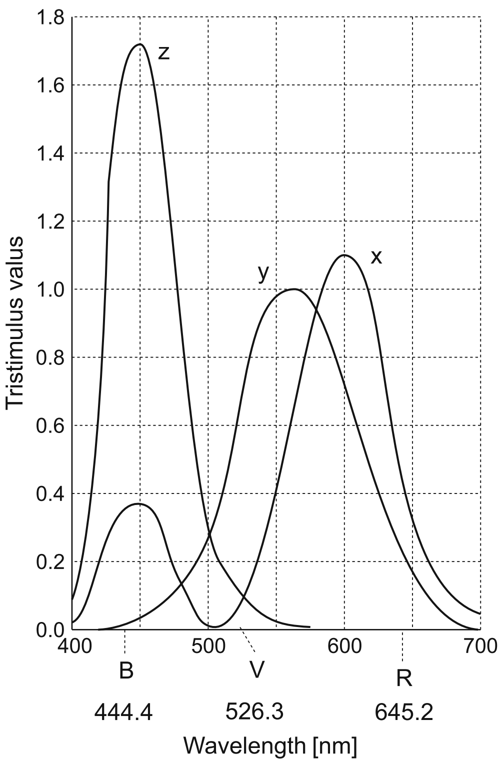


Fig. 4.10 Tristimulus functions defined by the CIE (*Commission Intemationale de l'Éclairage*, 1931) related to the three primary colours XYZ : RGB

4.3.3.1 Colour Definition

A colour can be quantitatively defined on the basis of the following elements:

- reflectance of the observed surface;
- curve of the energy source;
- three tristimulus functions of the eye;
- obtaining values that express, in numerical and absolute terms, the equivalent of the eye colour vision.

4.3.3.2 Chromaticity Coordinates

Since the sum of the values of tristimulus X, Y, Z represents a global effect of the colour, a term of normalized values is used as

$$
\begin{aligned}
\mathrm{x} &= \mathrm{X}/(\mathrm{X}+\mathrm{Y}+\mathrm{Z})\\
\mathrm{y} &= \mathrm{Y}/(\mathrm{X}+\mathrm{Y}+\mathrm{Z})\\
\mathrm{z} &= \mathrm{Z}/(\mathrm{X}+\mathrm{Y}+\mathrm{Z})
\end{aligned}
$$

resulting in

$$x+y+z=l \tag{4.10}$$

where x, y, and z have values ranging from 0 to 1. They are defined as chromatic coordinates or tri-chromatic coefficients and represent the normalization of each tristimulus value in comparison to the sum of the three, according to the CIE's System of Colorimetry X, Y, Z.

The normalization of tristimulus values reduces the number of variables. In fact, the two coordinates x, y are sufficient to calculate the value of z, which is the complementary value of the sum of x and y.

As modification and development of the triangle X, Y, Z, the CIE chromatic coordinates x, y assume a horseshoe shape (Fig. 4.11), where the place of the saturated colours, corresponding to the curvilinear segment, points out the wavelengths of the real colours. The coordinates of the luminous sources fall in the central zones. The standard source W has equivalent energy distribution of the three primary colours and coordinates x, y, z each with equal value of 0.333.

The linear segment, between blue, 400 nm, and red, 700 nm, is defined as axis of the magentas.

4.3.3.3 Hue, Saturation, Intensity

Three parameters are necessary to define a colour: Hue, Saturation and Intensity (HSI).

Hue is the name of a distinct colour of the spectrum, violet, blue, cyan, green, yellow, orange, red, and so on. It is the particular wavelength frequency. It regulates

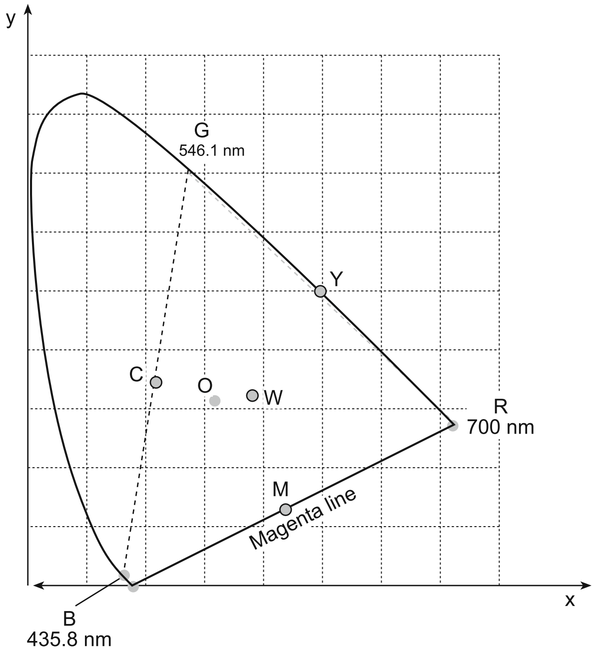


Fig. 4.11 Diagram of the chromaticity coordinates x, y representing the spectral and non-spectral colour distribution in the CIE classification system based on the imaginary primary colours X, Y, Z The triangle RGB shows the field of existence of all the possible colours starting from the three additive fundamental Red: 700 nm, Green: 546.1 nm, and Blue: 435.8 nm.
C: assay colour;
W: white with proportional content of RGB;
Y: Yellow, M: Magenta, C: Cyan.

the spectral association of a colour in the diagram of chromaticity. Hue is defined in the CIE diagram.

Saturation refers to the amount of white light, or greytone, mixed with the hue. The degree of saturation of a colour, or purity, is deduced from the saturated colour of reference. This characteristic is calculated as a percentage, on the basis of the ratio of the distance from the saturated colour of reference to the distance from the source.

Intensity, also called lightness or brightness, is the value of a colour. It is the amount of light or white it contains. The value of the colour *intensity* represents (as a percentage) the energy of the luminous source. It defines how bright the colour appears. When light is at its greatest intensity, colours will become bright; at its least intensity, colours become dim. The intensity is derived from the value of tristimulus Y. This implies that for reflecting surfaces (or transmitting) Y = 100%, while for absorbing surfaces Y = 0%.

4.3.3.4 CIE System

The CIE colorimetric system has the advantage of synthetically representing the spectral information coming from the observed objects through three numbers (by

X, Y and Z or by HSI); these terms of numbers need to be carefully analysed. It is essential to distinguish two concepts:

- the colour as a physical unit is independent of the visual process;
- the relativity of chromatic appearances: given the chromaticity and the intensity of a spectrum, it is not certain how the colour will appear if the conditions of observation are not known (illumination, contrast deriving from the contiguity with other colours, degree of fatigue of the eyes, topological circumstances of the observed object).

In conclusion, any measurement system, and consequently also any colour representation system, although precise, in perceptive terms will give our brain only an expectation of a given colour expressed in numerical terms, never the certainty.

4.3.4 Radiometric Terminology

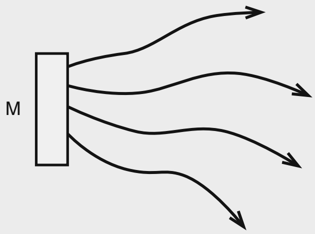
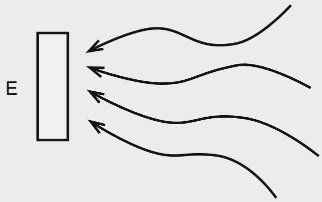
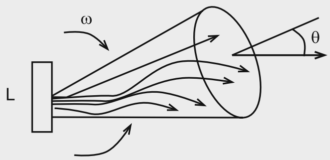
Radiometry is concerned with defining and measuring radiometric units resumed in Box 4.2, variable in function of the wavelength. Therefore when radiometric units are used, the interval of wavelengths to which they refer must be specified.

Since some of these refer to plane and solid angles, definitions of radian and steradian are recalled here:

- *radian* (rad): Standard International (SI) unit of plane angle equal to 180/π degrees.
- *steradian* (sr): SI unit of solid angle. There are 4π steradians in a sphere. The steradian, conical in shape, is the solid angle subtended at the centre of a sphere by an area on the surface of the sphere that is equal to the radius squared. The radian and steradian are dimensionless units.

Box 4.2 The most important radiometric terminology

Radiometric term	Symbol	Definition	Unit	
Radiant energy or electromagnetic energy	Q	Radiant energy is the energy of electromagnetic waves	Joule (J)	
Radiant flux or Radiant power	Φ	Radiant energy transferred from a point or from a surface to an other in the time	Watt (W)	

Radiometric term	Symbol	Definition	Unit	
Density of radiant flux		Radiant flux on a surface divided by the surface area Radiant exitance and irradiance: units and expressions are identical		
Radiant exitance	M	Radiant flux or radiation that is leaving an object or a surface	$W.m^{-2}$	
Irradiance	E	Density of radiant flux or radiation is incident on an object or a surface	$W.m^{-2}$	
Radiant intensity	I	Density of radiant flux emitted in a given direction from a surface per steradiant of solid angle of the receiver. Defined also as emittance	$W.sr^{-1}$	
Radiance	L	Radiance is a geometric radiation quantity that allows to describe how radiation is distributed in space. Radiation from a surface per unit area per steradiant of solid angle of the receiver, measured on a plane perpendicular to the considered direction.	$W.m^{-2}.sr^{-1}$	

Box 4.3 The most important radiometric terms used in remote sensing and correspondent units

Terms	Unit	Symbol (SI)
Length (l)	Meter	(m)
Wavelength (λ)	Micron	(μm) 10–6 m
	Nanometre	(nm) 10–9 m
Frequency (ν)	Hertz	(Hz)
(cycles per second)	Megahertz	(MHz) 106 Hz
	Gigahertz	(GHz) 109 Hz
Time (t)	Second	(s)
Mass (m)	kilogram	(kg)
Force (m ·l ·t–^{2})	Newton	(N)
Energy (m · l^{2} ·t –^{2})	Joule	(J)
Power (m ·l^{2} ·t–^{3})	Watt	(W)
Plane angle	Radiant	(rad)
Solid angle	Steradiant	(sr)

4.3.4.1 Radiance: The Radiometric Term Measured in Remote Sensing

The *radiance* (*L*) is the most important radiometric unit, as it describes what in the real world is measured by the sensors used in remote sensing (Box 4.3).

The optical-passive sensors on board satellite and aerial platforms, or field instrumentation, are sensitive to only one radiometric unit: the *radiance*.

The *radiance L* refers to the radiation according to a certain angle of observation, indicating the radiant outgoing flux per surface unit and per solid angle unit, measured on a perpendicular plane to the considered direction. The radiance is measured in Watts per square metre per steradian (W·m^{-2}·sr^{-1}).

$$L = \frac{1}{\cos\theta} \cdot \frac{\partial^2 \Phi}{\partial A \partial \Omega} \tag{4.11}$$

where θ : angle formed from the incident radiation to the plane; Φ : flux of the outgoing radiation; ∂A : infinitesimal portion of the plane reached by the incident radiation; $\partial \Omega$: infinitesimal portion of the solid angle.

Note: an infinitesimal or infinitely small number is a number that is smaller in absolute value than any positive real number.

The reason why the radiance has been chosen as a detected unit is related to its independence of the distance from the sensor to the detected surface. In fact, the radiance is the energy coming from a unitary surface per unit of solid angle (Fig. 4.12), and the width of the solid angle is independent from the distance of the point. Information is not lost in the space; just the sensor surface must be progressively increased going further from the surface.

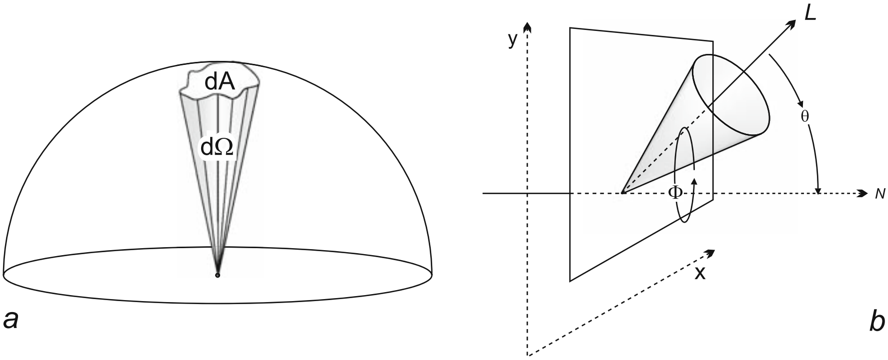


Fig. 4.12 (**a**) Solid angle expressed in steradian (sr); (**b**) Radiance L (W m–2 sr–1): radiation depending on the observation angle, measured on a plane perpendicular to the considered direction N

Also the radiance varies in function of the interval of wavelength, and it must be specified each time.

The *reflectivity* ρ is the ratio of reflected energy to incident energy: E_r/E_i expressed in values ranging from 0 to 1 (in percent, better known as *albedo*) is frequently used sometimes incorrectly instead of the radiance, even if it is not a radiometric unit. When reflectivity is measured referring to a definite wavelength interval, it must be indicated as spectral reflectance. The width of the angle at which measurements of reflectance is carried out should also be specified.

Box 4.4 lists the considered radiometric units and the associated photometric units.

Box 4.4 The most important radiometric terms, corresponding photometric terms and units

Radiometric term	Unit	Symbol	Photometric term
Radiant energy	Q	(J)	Luminous energy
Radiant flux	Φ	(W)	Luminous flux
Density of radiant flux			Density of luminous flux
Radiant exitance	M	$(W \cdot m^{-2})$	Luminous exitance
Irradiance	E	$(W \cdot m^{-2})$	Illuminance
Radiant intensity	I	$(W \cdot sr^{-1})$	Luminous intensity (luminance, brightness)
Radiance	L	$(W \cdot m^{-2} \cdot sr^{-1})$	Radiance

Photometry is the branch of applied optics studying the methods for measuring the characteristics of a luminous radiation beam in the visible. Photometric units

measure the light emitted by a source or received by a body. A photometric unit is related to each radiometric unit; the main units of photometric quantities are

- *lumen* (lm) defined as the luminous flux emitted by a candela in the solid angle of a steradian;
- *candela* (cd) defined as 1/60 of the luminous intensity emitted by a blackbody, or integral radiator, at platinum solidification temperature, in normal direction to the exit hole having a section of 1 cm^2; it is the main measurement unit of luminous intensity.

4.3.5 Spectral Response

The visualization of satellite digital images is the result of the conversion of the signal recorded by the sensor into discrete value or Digital Number (DN) associated to each elementary pixel cell. Since this signal is a function of the physical–structural characteristics of the detected object, it is possible to establish a correspondence between the quantity and the quality of the reflected energy and the nature of the objects, in relation to the different wavelengths.

The signal recorded by the sensor can be represented in graphical terms as reflection capability in function of the wavelength: the *spectral response*, a key instrument for satellite image quantitative analysis (Fig. 4.13). (See my note related to the term: spectral response.)

When the electromagnetic energy emitted by the Sun reaches the surface of an opaque body, it is partially absorbed and partially reflected. The reflection can be

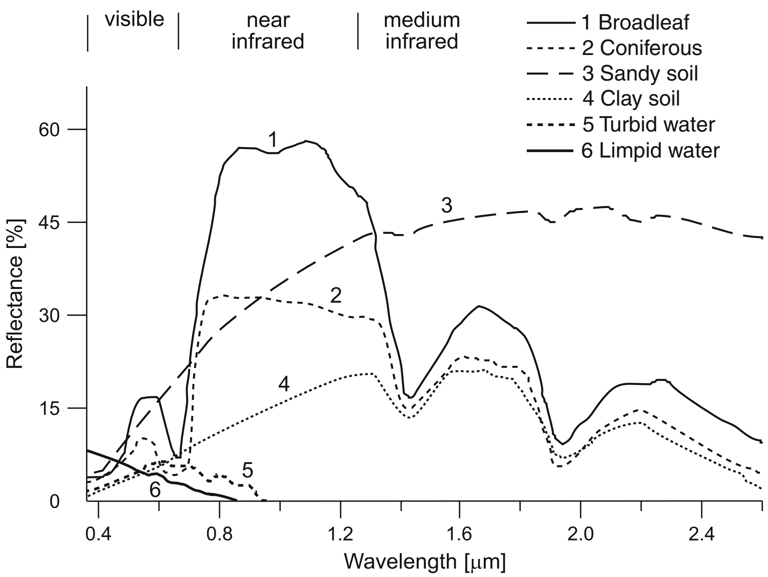


Fig. 4.13 Spectral responses of some natural terrestrial elements

specular or diffuse in relation to the object or to the angle of view (see Fig. 4.7). The energy reflected in the visible contains spectral information also concerning the colour of the reflecting surface. The reflectance is determined by the geometric structure of the surfaces, by the bodies' nature, by vegetation pigmentation, etc. For example, chlorophyll strongly absorbs the radiant energy in wavelength intervals near 0.45 μm (blue) and 0.65 μm (red), while reflecting the green radiation, near 0.55 μm wavelength, visually perceived as the leaves' colour and reflects more in the region of near infrared from 0.8 to 1.3 μm.

Passive remote sensing in optical bands from ultraviolet (0.3 μm) to medium infrared (2.5 μm) deals with studying and measuring the characteristics of reflection, aiming to identify *iso-behaviour* surfaces that should correspond to similar objects.

Anyway the spectral response is not univocal; the spectral behaviours of objects belonging to the same class can generate different spectral responses.

The factors that produce variations in the spectral reflectance curves can be

- *static*: such as slope and ground exposition;
- *dynamic*: inducing differences in spectral behaviour of the same elementary cell over time. Among them are vegetation phenological phase, phytosanitary conditions and fractional cover, soil surface humidity, atmospheric transparency, Sun position, etc.

The Sun elevation angle causes differences in the reflectance, also in relation to the soil. The reflectance of a sandy soil is more sensitive to illumination variations than the reflectance of vegetated soils. This is mainly due to geometrical aspects for which sandy soils can be considered more similar to a smooth surface than a surface covered by tree canopies.

The curves of spectral reflectance of an object change in relation to the variability of local environmental conditions and the sensors collect the energy reflected not only from a single cell but also from the nearest neighbours, in terms of direction and instant of observation.

In case of mixed pixels, the recognition of a spectral response can be very difficult.

Stressed vegetation shows reductions of absorption in the medium infrared and green intervals due to the decrease of water content and pigments. Figure 4.14 shows the role of chromatic characteristics of some minerals: the hematite, an iron oxide, has a peak in the red, while the biotite, black-coloured, has low reflectivity over the entire spectrum; the albite, white-coloured, is the most reflecting among the selected minerals.

4.3.6 Electromagnetic Radiation–Atmosphere Interaction

The word atmosphere refers to the gas layers, as a whole, surrounding the Earth: its complex chemical, physical, optical and dynamical processes affect lots of phenomena occurring on the Earth's surface.

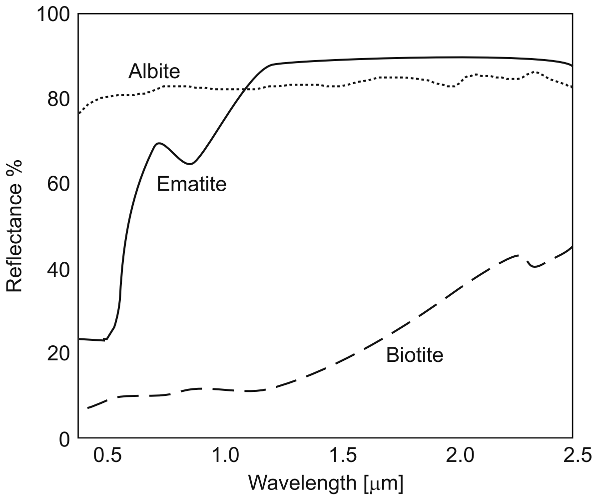
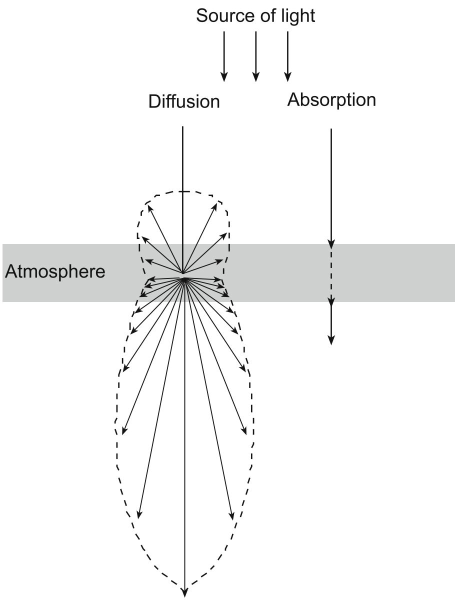


Fig. 4.14 Spectral responses of some minerals

Fig. 4.15 Absorption and diffusion of radiation



The interest in this interaction is related to the fact that atmospheric components diffuse, refract, reflect, absorb and emit electromagnetic radiation changing the original radiance of the objects observed by a remote sensor (Fig. 4.15).

Atmospheric effects depend on the wavelength and are both additive and multiplicative.

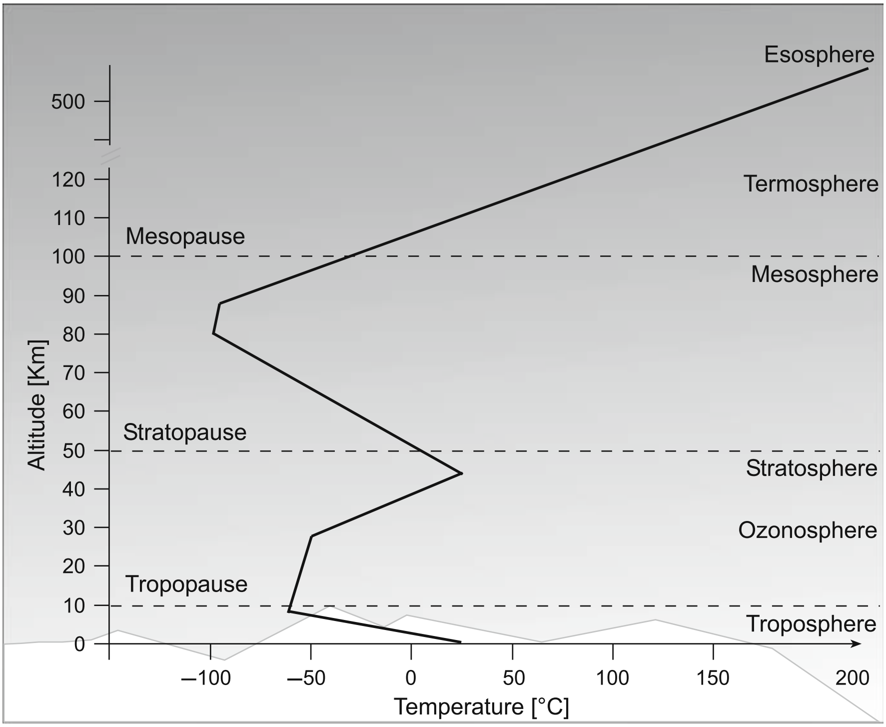


Fig. 4.16 Temperature values of the atmosphere stratification at different altitude

The atmosphere extends in height about 500 km above sea level, including the thermosphere and shading beyond the exosphere in the interplanetary vacuum. It is subdivided as (Fig. 4.16):

- Troposphere (from 0 to 10–12 km), currents and perturbations form in its lower part, where the temperature decreases with increasing altitude.
- Stratosphere (from 12 to 50 km), where temperature rise from very low (−60 °C) to +10 °C at the upper limit (stratopause) with the next layer.
- Ionosphere extends from 50 to 6370 km, where ions and electrons are so abundant as to modify the propagation of electromagnetic waves.

Atmospheric layer transparency to the transmission of electromagnetic energy varies in the spectrum and, also where it is more transparent, transmittance values can rapidly change varying the meteorological conditions altering the composition of the atmosphere (Fig. 4.17).

The constituents of the atmosphere can be divided into two groups:

- *permanent constituents*: nitrogen (78%), oxygen (21%), argon (0.93%) and carbon dioxide (0.03%);

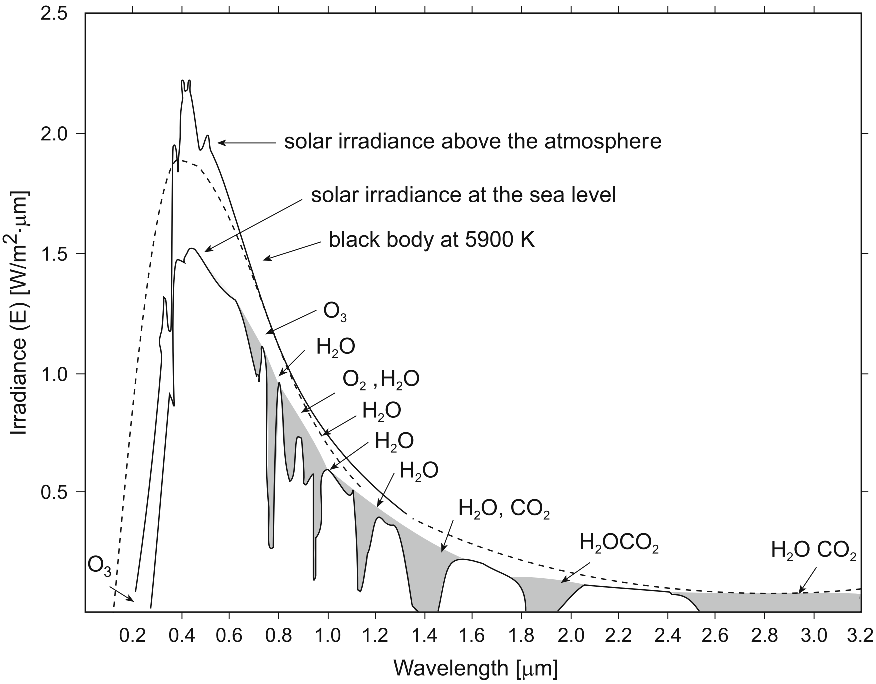


Fig. 4.17 Curves of Sun irradiance outside the atmosphere and at sea level compared with the blackbody curve at 5900 K, corresponding to the Sun curve

- *variable constituents*: ozone, water vapour and other gases especially in the industrial areas.

4.3.6.1 Atmospheric Effects in the Visible

Atmospheric absorption occurs when the atmosphere stops or reduces radiation transmission. The atmosphere gains some energy, which is reemitted at longer wavelengths. Quantity and distribution of permanent constituents are known, evaluating their effect on the transmission of energy, while variable atmospheric aerosol is time and place of observation dependent. The effect of the aerosol on the propagation of the electromagnetic wave is a function of physical characteristics like shape, spatial distribution and refraction index, which depends from the chemical composition.

Clouds and fog make the evaluation of the direct radiation attenuation and the anisotropic (non-constant directional intensity) scattered radiation difficult. Image data availability are compromised too.

Atmospheric gases, aerosols and vapours contribute to absorb, diffuse and refract the direct solar radiation and the radiation reflected by the surface:

- spectral *absorption* reduces the quantity of energy available in a given wavelength;
- *scattering* distributes the radiation modifying its propagation direction;
- *refraction* modifies the radiation path direction.

Moreover:

- the molecular structure of each chemical component has specific bands of absorption, that is, different gases absorb in different wavelength intervals $\Delta\lambda$ ($\lambda_1 - \lambda_2$);
- scattering depends on molecular concentration and is expressed through a function of the atmospheric density and pressure;
- refraction depends on the optical characteristics of the media density.

These phenomena cause a decrease of contrast among the objects in the image (Fig. 4.18).

Atmospheric absorption and scattering have a large influence on the ratio (recorded radiance/original [reflected or emitted] radiance). The average annual flux density of the incident radiation on the terrestrial atmosphere is 341.8 W/m^2, and, due to atmospheric attenuation processes, only 147.3 W/m^2 reach the Earth's

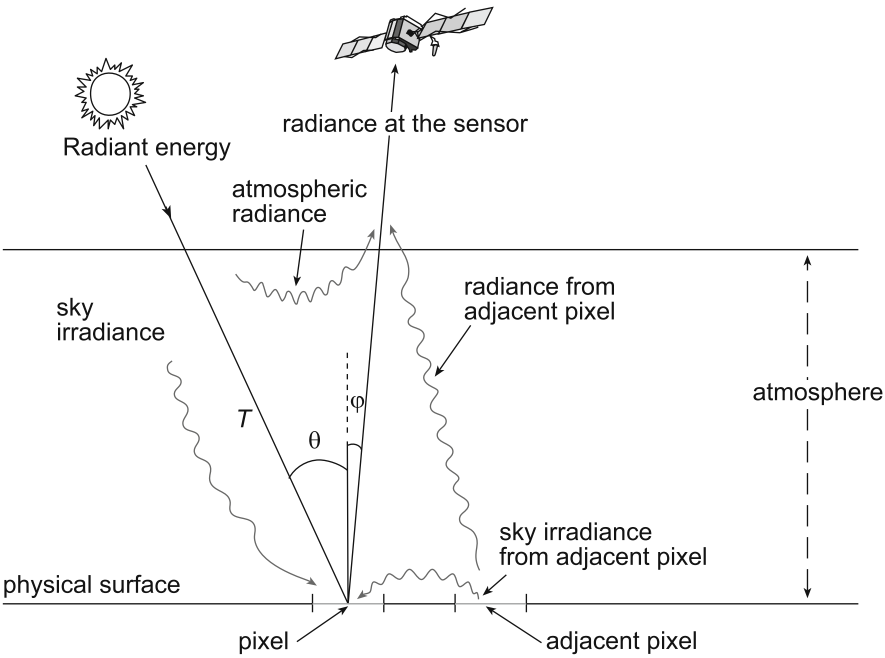


Fig. 4.18 Atmospheric effect on measuring the radiance of a pixel in a passive remote sensing system. Energy reflected by a pixel and measured by a remote sensor derives from a complex interaction of phenomena on the ground and in the atmosphere
θ: Sun zenith angle, referred to the perpendicular of the Earth's surface. φ: sensor view angle, referred to the perpendicular of the Earth's surface

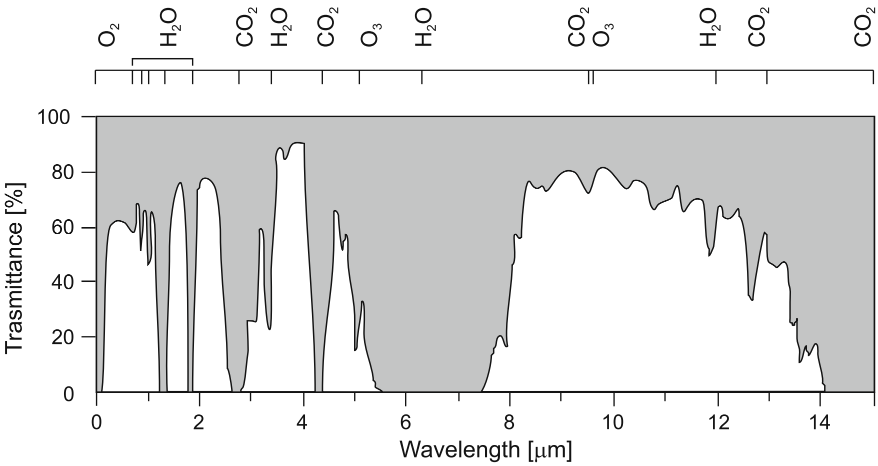


Fig. 4.19 Atmospheric windows of electromagnetic radiation transmission referred to the gas absorption. The white area represents the atmospheric windows, where the radiation passes through

surface every year. Moreover absorption is very selective: each atmospheric component (aerosols and chemical species in general) absorbs the radiation in function of its wavelength.

The chemical compounds responsible for the absorption of X and ultraviolet (UV) radiations are O, O_2 and O_3, while H_2O and O_3 absorb in the visible range, and H_2O and CO_2 in the infrared. As a consequence, atmosphere transmissivity τ (or transparency) of radiation is variable, limiting the possibility to collect signals in the EM intervals *atmospheric windows* (Fig. 4.19). Transmissivity also depends on the time and place of acquisition. The most used atmospheric windows are 1, 2, 3 and 6, indicated in Table 4.4, where transparency is highest and the effects of reflection and emission are well distinguished.

Scattering has a significant role at altitudes lower than 50 km, where aerosols and air molecules cause it. The effects are visible to the naked eye: the red colour of the

Table 4.4 Atmospheric windows, intervals in the electromagnetic spectrum, where the trasmissivity enables the remote sensing application

Atmospheric window	Spectral range (μm)		
1	0.3–1.3	UV, visible, near IR	Reflected
2	1.5–1.8	Medium IR	Reflected
3	2.0–2.6	Medium IR	Reflected
4	3.0–3.6	Medium IR	Reflected/emitted
5	4.2–5.0	Medium IR	Reflected/emitted
6	7.0–15.0	Thermal IR	Emitted

Sun on the horizon, the blue colour of the sky (when clear) or white (when rich in suspended particles) is phenomenon caused by the scattering.

The atmospheric diffusion, or scattering, phenomenon consists in the multiple re-directing of the electromagnetic energy by suspended particles or gas molecules. The radiation scattering intensity depends on

- gas particles size;
- number of gas particles per unit of volume;
- thickness of the atmosphere;
- radiation wavelength.

Rayleigh scattering occurs when atmosphere particles have a diameter ϕ much smaller than the wavelength of the incident radiation ($\lambda >> \phi$).). These particles can be dust particles or common atmospheric gas molecules like N_2 and O_2. Rayleigh scattering depends on the wavelength: shorter wavelengths undergo a higher scattering. Scattered radiation intensity is inversely proportional to the fourth power of the wavelength (*scattering*: $1/\lambda^4$). Blue light is scattered four times more than red light. This scattering is responsible for the blue colour of the sky and the red-orange colours at sunset.

Mie scattering is caused by atmospheric molecules such as dust, smoke, pollens and water drops. These particles have a diameter ϕ approximately equivalent to the radiation wavelength ($\lambda \cong \cong \phi$). Also Mie scattering depends on the wavelength, but not in a simple way like Rayleigh.

Non-selective scattering occurs when the particles, water drops or dust, are much bigger than the wavelength ($\lambda << \phi$). This kind of scattering, independent of the wavelength, is responsible for the white colour of fog and clouds, as different wavelength radiations are scattered in the same way.

Refraction is the deviation of the electromagnetic radiation path from the original direction, occurring in correspondence of the separation surface between two optically different media when the radiation goes from the first to the second medium. This phenomenon occurs also in the atmosphere when the light passes through different density atmospheric layers.

4.3.6.2 Atmospheric Effects in the Thermal Infrared

In the spectral range of thermal infrared, the radiation emitted by the Earth's surface is much higher than the reflected energy, therefore in this region of the electromagnetic spectrum, it is possible to relate the radiance recorded by a remote sensor with the temperature of the investigated surface.

As well as in the spectral interval of the visible, a procedure of radiometric calibration and atmospheric correction needs to be carried out before using data collected by remote sensors measuring the radiation emitted by the Earth's surface.

Absorption of thermal infrared radiation is mainly due to water vapour, then to carbon dioxide and ozone (Table 4.5). Water vapour is in the lower layer of the atmosphere, the troposphere, at altitudes lower than 10 km. Its presence can vary

Table 4.5 Molecules in the atmosphere responsible for radiation absorption in the medium and far infrared, and central wavelength of the corresponding absorption windows

Component	Molecule	Absorption band (μm)		
Water vapour	H_2O	2.7	3.2	6.3
Carbon monoxide	CO	4.8		
Carbon dioxide	CO_2	2.7	4.3	15
Ozone	0_3	4.8	9.55	14.2
Nitric oxide	NO	4.7	7.8	
Methane	CH_4	3.2	7.8	

in relation to latitude and season, and its concentration in the air quickly decreases with increasing altitude. Carbon dioxide, a permanent gas, is considered constant and uniformly distributed in the atmosphere; carbon dioxide bands of absorption are centred respectively around 2.7 and 15 μm (Fig. 4.20).

Ozone absorption is considered only from space. The thickness of the ozone layer, extended from 20 to 30 km of altitude, varies during the day, reaching the highest values when the sunlight is strongest: the sunlight, catalyzing the interaction between the ultraviolet radiation and oxygen, supports the formation of ozone.

Besides absorption due to atmospheric gases, infrared radiation can be refracted by the same gas molecules and by other particles in the atmosphere, like water drops and ice crystals coming from fog, clouds and rain, or carbon particles produced by combustion.

If in the visible the scattering phenomena are dominant, in the thermal infrared with clear sky the process that causes most problems is atmospheric absorption. Nevertheless in the presence of overcast sky or fog, the phenomenon of scattering becomes relevant also in thermal infrared.

Therefore the atmosphere limits the observation of the terrestrial surface to the wavelengths of thermal infrared, where atmospheric absorption phenomena are minor (atmospheric windows). These wavelengths correspond to two intervals of the spectrum: 3.5–5 and 8–14 μm. In the case of low flight observations, the atmospheric windows are more extended due to less opacity.

The two atmospheric windows are distinguished since the radiation emitted and reflected energy is present in daylight (3.5–5 μm), whereas in the 8–14 μm range only the emitted radiance can be detected. The atmospheric window ranging in 3.7 μm can be used by night acquisitions, excluding reflection phenomena.

The window ranging in 11 μm is preferred for temperature measurement for the following reasons:

- the terrestrial surface has its maximum radiation emission at 11 μm (according to Wien's displacement law);
- the reflected component is negligible allowing acquisitions in daylight.

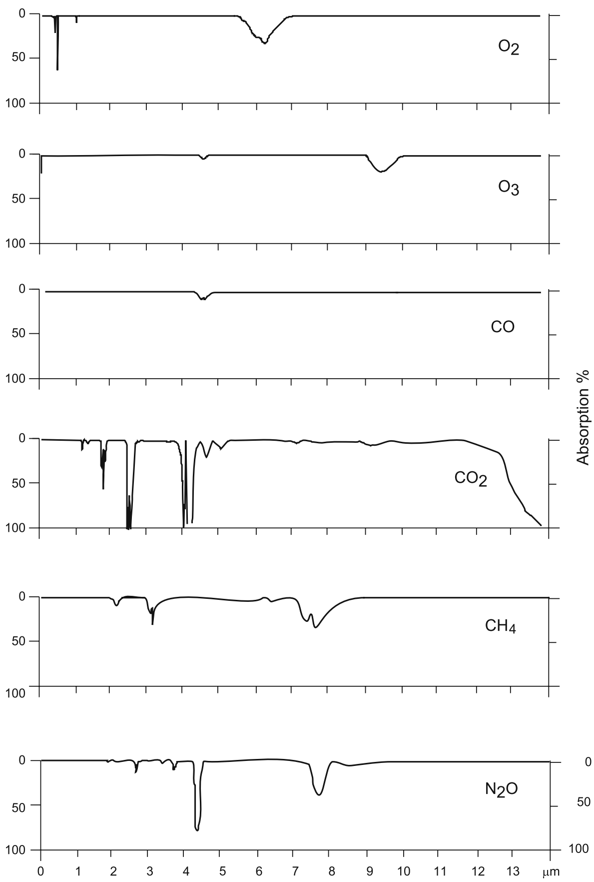


Fig. 4.20 Spectral absorption diagrams of several atmospheric elements: H_2O: water; O_2 and O_3: oxygen and ozone; CO and CO_2: carbon monoxide and dioxide; CH_4: methane; N_2O: nitrogen oxide

4.4 Active Remote Sensing in the Microwave

The first studies of radar, a typical microwave instrument, began at the same time in the United States and Europe in the 1930s. The approaching Second World War quickened and kept secret the research, which was carried out in countries involved in the war and mainly focused on military purposes, as often happens in these circumstances.

These studies were more advanced in Great Britain, United States and Germany; Great Britain was protected by a series of radars intercepting German bombers and as matter of fact this tool played a determining role in the outcome of that conflict.

Only more recently has active remote sensing been used in civil applications and in particular in Earth Observation (EO). The first radar instrument for this use was

launched on the Satellite Seasat in 1978. It provided data and images of the Earth for 105 days and opened the way to the next missions by many Spatial Agencies: European (ERS-1 and ERS-2, Envisat), Canadian (Radarsat-1 and 2), Japanese (Jers-1, Alos), German (TerraSAR-X) and Italian (COSMO/SkyMed constellation) (see also Chapter 6).

Radar is an active sensor made by a sensing system based on the principle of echo, in which the transmitter periodically radiates energy in the form of high power and very short duration microwave impulses towards the observed scene. The radar system coherently records the backscattered signal and derives information about the distance of many backscatterers in the scene by calculating the temporal delay from sent signal to received echo.

A strongly directive antenna irradiates radar impulses by transmitting them towards the target at flight speed.

If transmitted impulses do not encounter any obstacles, they are not backscattered, while if they encounter an object, a ship, a mountain, a building, a small part of irradiated energy goes back to the transmitting antenna in a very short time in the form of an echo and is visualized as light spots.

As the propagation speed is exactly known, the distance of the target can be retrieved by the time the signal takes to reach it and go back.

Independently of the architecture used, a classic radar system can perfectly retrieve accurate information about the distance and speed of isolated objects, in any weather and illumination conditions. Therefore it is widely used in sighting flying planes and sailing ships. In those applications, as well as EO for different purposes, one is aiming to study a complex scene where it is desirable to distinguish as many details as possible. The system capacity of placing different objects exactly in the space and distinguishing near objects related to its geometric resolution is important. The radar works in a time domain, since space can be retrieved, the light speed being constant.

Working with active microwaves has many advantages mainly related to the capacity to cross through the cloud blanket (in fact atmosphere attenuation is in practice null for wavelengths higher than 3 cm) and to the independence of the Sun illumination. This permits, for instance, the surveying of tropical areas, often covered by clouds and strongly influenced by air humidity. A radar system comprises a transmitter emitting a beam of electromagnetic waves, usually modulated by pulses and with a depression angle, and by a receiver measuring the intensity of the backscattered radiation, called *backscattering radiation* (Figs. 4.21 and 4.22). Usually the same antenna is used to alternatively transmit and receive (Fig. 4.23).

Applications of remote sensing in radar bands are developing and in some cases already offer interesting results. Radar is used for structural geology and hydrogeology surveys, in research about the sea's surface and studies about carsism, volcanic, alluvial and glacial morphology, for applications in agriculture and forestry and for monitoring catastrophic events.

Radar images are made up of pixels whose value is proportional to the power of the signal reflected from the corresponding cell on the ground to the antenna. The

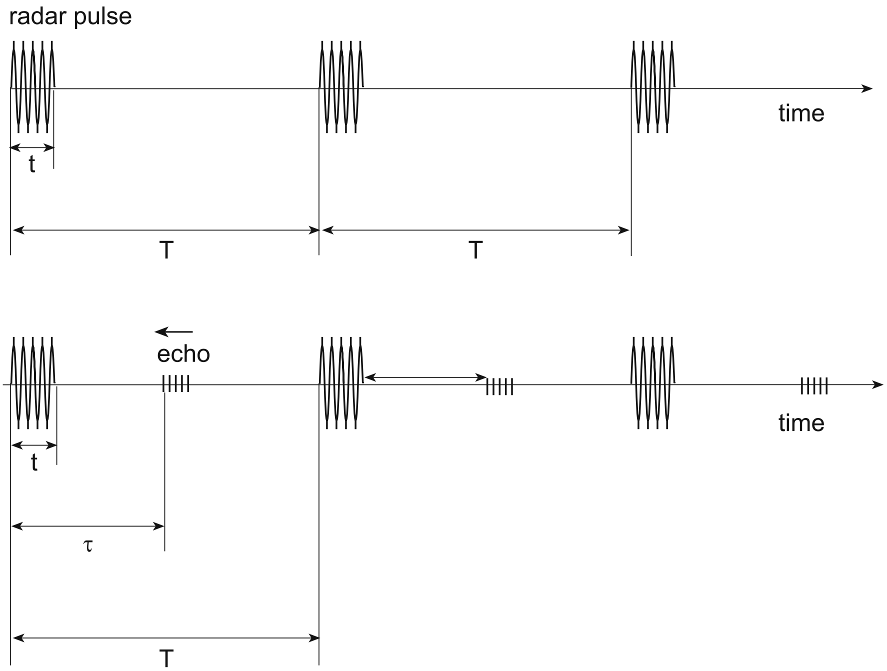


Fig. 4.21 An active microwave sensor (imaging radar carried by either aircraft or satellite) is an instrument that transmits a microwave signal and then receives its reflection as the basis for forming digital or pictorial images of the Earth's surface

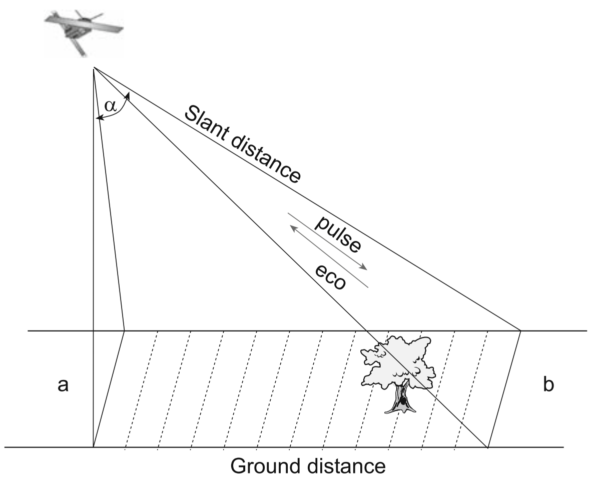


Fig. 4.22 Geometric differences of the optical and radar systems: (**a**) geometry of a vertical acquisition (nadir) of an optical image and (**b**) geometry of the acquisition of a radar image; α: incident angle of acquisition of the radar image

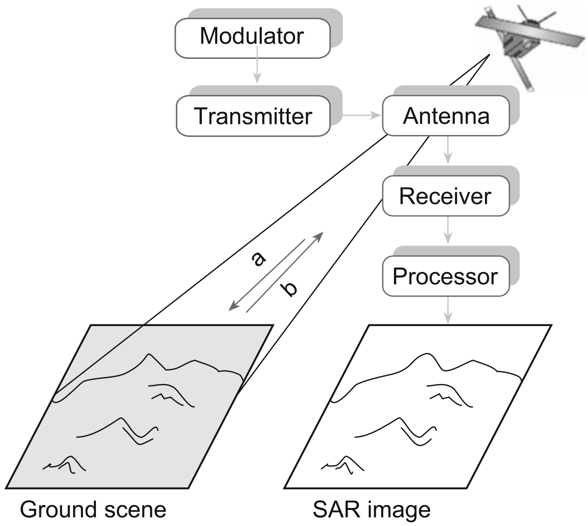


Fig. 4.23 The fundamental principle of radar imagery acquisition: (**a**) radar pulse and (**b**) echo or backscattering In the radar (radio detection and ranging), the ranging capability is achieved by measuring the time delay from the time a signal is transmitted towards the terrain until its echo is received. The distance is a function of the propagation velocity of the signal.

equation for radar bands retrieving the average backscattered energy per each image point is

$$< P_r >= P_t \left[\frac{G^2 \lambda^2}{(4\pi)^3 R^4} \right] \sigma^0$$

where $< P_r >$: average received power from the antenna; P_t: transmitted power from the antenna; G : gain for the radar antenna; λ: radar wavelength; R : range from the radar antenna to the target; σ^0 : coefficient of backscattering (targets scattering cross section).

The backscattering coefficient (σ^0) to which the scattering by the surface of the radar bands is related depends on the following parameters as a function of the radar acquisition phase (Fig. 4.24):

- *terrestrial surface objects parameters*: roughness, water content or humidity, dielectric constant, electric resistivity, surface geometry, surface position;
- *incident signal wavelength (λ)*: λ theoretical interval that can be used for radar observation ranges from 1 mm to 1 m, as already said, and it is in practice used from 0.8 cm to 1 m (Table 4.6). Since a body must have a size similar to the electromagnetic wavelength to be able to interact with it, atmosphere components are not obstacles to radar wave penetration;
- *signal polarization*: a radar signal can be generated in such a way so that wave oscillations are limited to only one plane perpendicular to the propagation direction: in this case the wave is said to be polarized (Fig. 4.25). Independent of the wavelength, it can be transmitted or perceived in different polarizations: usually vertical (V) and horizontal (H) planes are considered. Four combinations are possible:

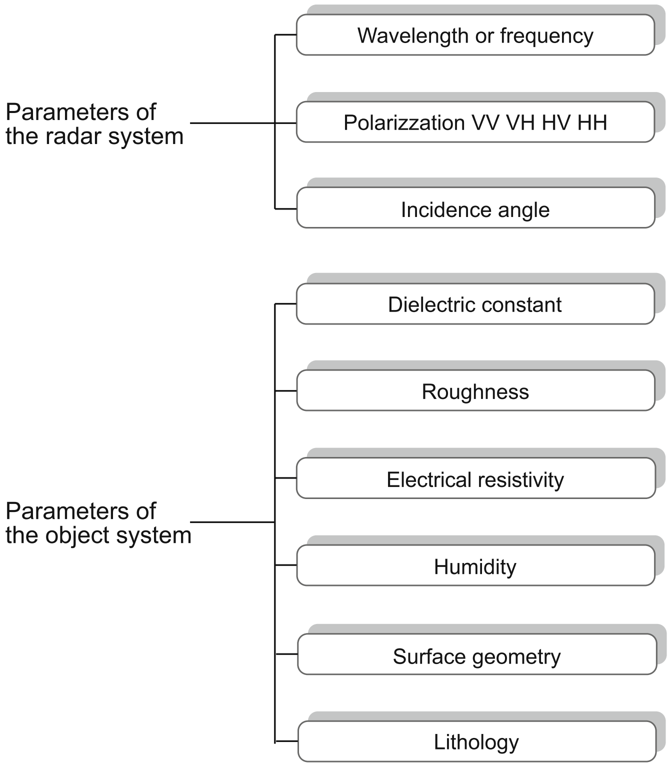


Fig. 4.24 Parameters of the radar system and of the objects on the Earth's surface that could influence the backscattering

- HH and VV, waves transmitted and received respectively in horizontal and vertical polarization;
- HV and VH, waves transmitted and received with cross-polarization.

The acquisitions in radar bands allow one to control the radiation emitted by the antenna in:

- energy;
- frequency (MHz);
- polarization (HH, HV, VH, VV);
- direction (angles of depression and azimuth) (Fig. 4.26)

Table 4.6 Radar band defined by wavelength and frequency: the most used intervals are in band X, C and L

Band	Wavelength (cm)	Frequency (GHz)
K	0.83–2.75	36.0–10.8
X	2.75–5.21	10.9–5.74
C	5.22–7.14	5.75–4.20
S	7.15–19.74	4.21–1.54
L	19.75–76.9	1.55–0.39
P	>76.9	<0.39

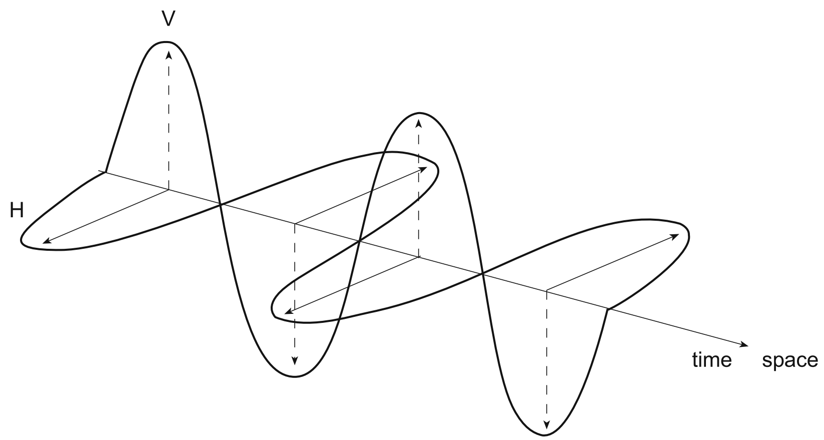


Fig. 4.25 Polarization of the radar frequency transmitted by the antenna; V: vertical; H: horizontal. If the backscattering signal is recorded in V polarization, the resulting polarization signals are respectively VV and HV. Inversely, with H polarization reception, polarizations VH and HH are obtained

The different shapes of objects can influence the reflected wave polarization, which then provides information about their identity. Moreover the type of polarization influences also the radar signal penetration into canopies with vertical structures and those planted in strips, like different types of crops (see Plate 4.6);

- *inclination angle of the acquisition*: the reflected energy is influenced by inclination angle variations. Higher inclination angles allow a wider ground exploration and a better image resolution, although the reflected signal (backscattering) intensity decreases; in the case of satellite sensors, incidence angle variations can be disregarded.

They also provide information about sub-surface phenomena, as a function of the biomass density and water content, using parameters describing surface properties different to the ones described by the optical bands. Radar signal penetration depth increases as the wavelength increases and is a function of the characteristics of the detected surface. The resulting intensity is related to the quantity of energy backscattered by the elements present (*scatterers*).

Radar image geometry is derived by the manner of data acquisition, as shown in the simplified scheme in Fig. 4.23: the radar system sees the ground with a perspective corresponding to an oblique line (*slant range*) from the antenna to the object on the ground. The back signals, echo radar, of the emitted impulses are registered in a temporal sequence and provide the image transversal component (*across track*). The longitudinal component (*along track*) results from the continuous emission of impulses while the system including the antenna on the platform moves along the flight direction. If the resolution along the flight line, or azimuth resolution, is inversely proportional to the dimension of the antenna sending the signals

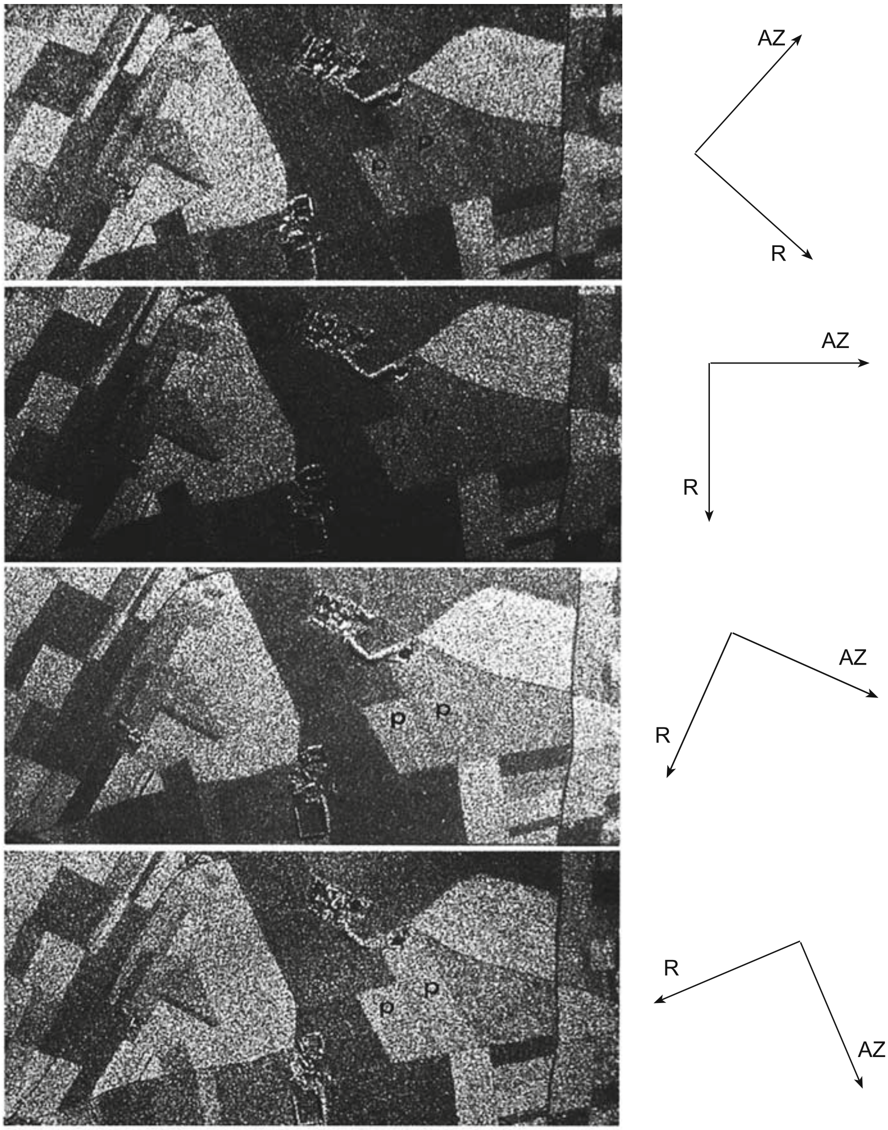


Fig. 4.26 Airborne SAR images registered from different view angles and azimuth. R: view perspective; AZ: azimuth direction; P: cultivated field of potatoes

and proportional to the radar-target distance and to the radar wavelength, the spatial resolution in distance ρ (*slant range*) is proportional to the transmitted impulses duration (τ):

$$\rho = c.\tau/2 \tag{4.13}$$

where c is the light speed.

To obtain distance resolutions lower than 10 m, it is necessary to transmit impulses with duration lower than 66 ns or to use a frequency higher than 15 MHz. In practice linearly modulated at much higher duration frequency impulses are transmitted, and then compressed with a filter adapted for data numerical processing.

With regard to the azimuth direction, the movement of the platform in relation to ground objects is used to synthesize a large-size antenna while in reality, it is small (from which the name of these *Synthetic Aperture Radar* systems was derived, abbreviated as SAR). In fact to have a resolution of 5 m at the frequency of 1 GHz and altitude of 800 km, a *real* antenna longer than 40 km would be required. With the SAR, the equivalent antenna is about 10 m long.

4.4.1 Radar Versus Optical Systems

According to the previous descriptions concerning passive optical and active radar, remote sensing shows different systems having conceptual and instrumental characteristics specific for observation from ultraviolet to microwaves (Box 4.5). Radar can provide different and more continuous information compared to optical sensors, and in particular it presents the following advantages:

Box 4.5 Satellite remote sensing from the visible to the radar

Characteristics	Visible/near IR	Thermal infrared	Synthetic aperture radar (SAR)
Wavelength(λ)	0.4–2.2 μm	10–12.5 μm	3–60 cm
Frequency	–	–	–
Resolution	0.61–1100 m	60–120 m	10–30 m
Type of sensor	passive	passive	active
Radiation source	Sun	Blackbody	Radar
Geophysics parameters	Albedo	Temperature	Dielectric and geometric properties
Through the clouds	No	No	Yes
Penetration in			
- soil	No	No	Yes
- vegetation	No	No	Yes
- water	Yes (visible)	No	No
Sun independent	No	Yes	Yes
Salt and pepper effect (speckle)	Low	Low	High
Geometric effect	(No)	(No)	Yes
Operative since	1972	1982	1991

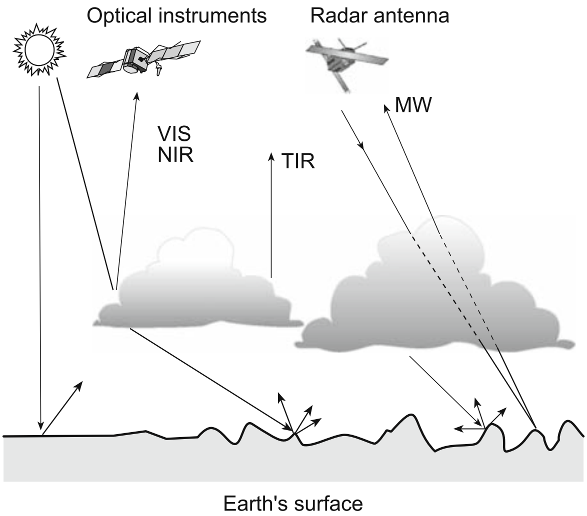


Fig. 4.27 The behaviour of the radiation at different wavelengths such as visible (VIS), near infrared (NIR) and thermal infrared (TIR) in the presence of clouds are not transmitted. In the radar band (MW), the signal overcomes that obstacle and reaches the Earth's surface, anytime (day and night)

- *any-time capacity*: it is insensible to the cloud cover, and therefore allows undisturbed multi-temporal detection. This is very important where it is difficult to obtain images through optical sensors requiring clear sky conditions (Fig. 4.27);
- *complementarity*: radar usually describes shape and structure of the objects and also provides data about soil moisture content and vegetation, while optical systems are sensitive to colour and temperature of the objects (Box 4.6). The information collected by the two kinds of sensors can be considered complementary (Figs. 4.28 and 4.29);
- *penetration*: radar waves, especially the longest (L 23 cm and C 5.6 cm bands), can pass through biomass and beneath it, collecting information about soil moisture (Fig. 4.30);
- *relief sensibility*: radar with lateral vision, which produces images, obliquely irradiates a ground portion enhancing morphological aspects such as roughness and slope. This sensitivity appears in the image as shadows characterizing mountainous or wrinkled areas; in Figs. 4.31 and 4.32, a smooth surface specularly reflects the incident radiation. The reflected radiation follows a path external to the antenna domain, and the corresponding areas will look dark in the image (low Digital Numbers DN near zero). The opposite happens in the case of a rough surface (DN tend to 255, highest values in 8-bit images);
- *coherent illumination*: this system generates images characterized by intensity and phase (whose usefulness is discussed in Section 4.4.3.1).

Box 4.6 Synergisms among some acquisition systems: optical-passive reflected and emitted, and radar-active

Satellite sensor	Strength	Spectral range	Measure
SPOT-HRV IRS-LISS4 Landsat7-ETM+ Pan	Medium geometric resolution (5–10 m)	Visible (VIS) Near infrared (NIR)	Surface reflectivity
Landsat7-ETM+ EOS/AM-Modis, Aster	High spectral resolution (6–15 bands)	Visible (VIS) Near infrared (NIR) Thermal infrared (TIR)	Surface temperature
ERS-SAR Envisat-ASAR	Penetration of clouds, soil and vegetation	Active microwaves	Surface roughness Soil moisture

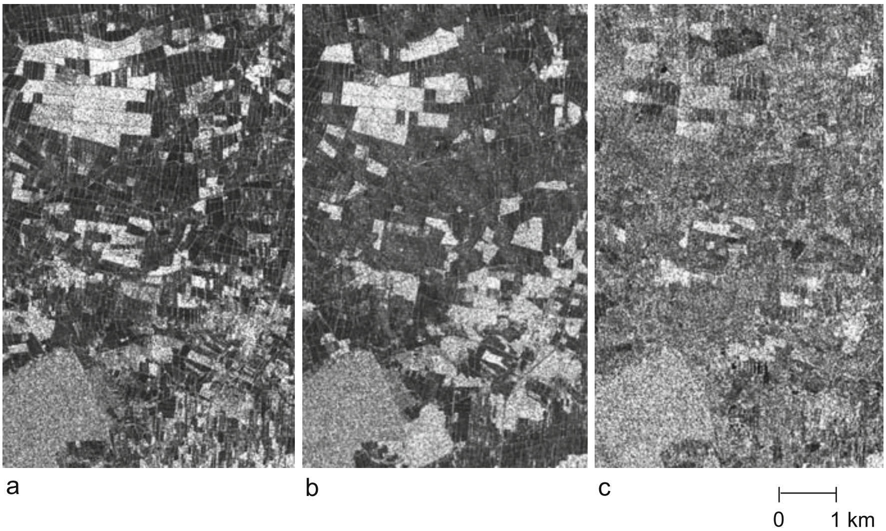


Fig. 4.28 The radar image is represented by the structure and texture influenced by the ground pattern and by the speckle, the basic noise of the signal. ERS-1 SAR images of (**a**) April, (**b**) May and (**c**) July showing the increment of the noise associated with the biomass growth. The noise increases the salt–pepper effect (*speckle*) in the images

4.4.2 Radar Systems

Geometric resolution of a real antenna depends on its size in relation to the used radiation wavelength. Since the frequencies commonly used in radar systems have wavelengths ranging from 0.1 to 1 m, an increase in geometric resolution implies the building of much bigger antennae. This solution is unfeasible for many applications

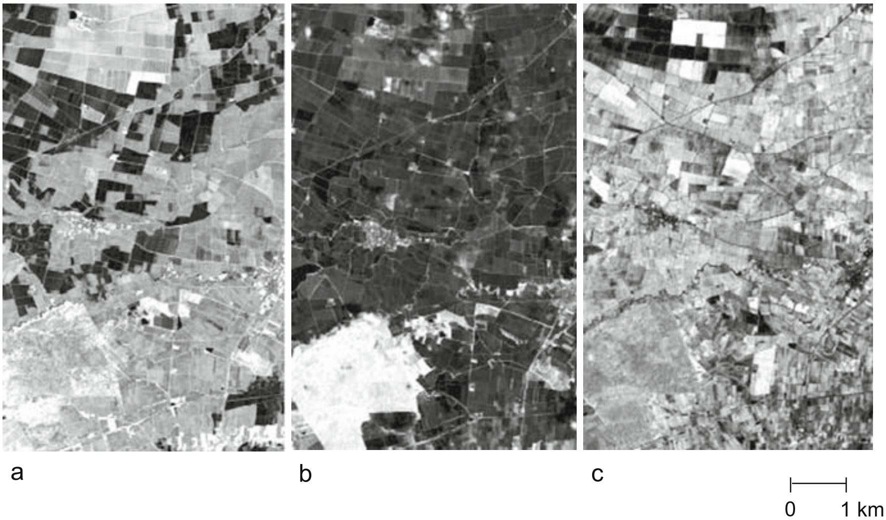


Fig. 4.29 Landsat optical images, in the near infrared in (**a**) April, May and July, comparable with the radar images in Fig. 4.28, in a rice field, are in the Vercelli Province, northern Italy. In (**b**) the *Partecipanza wood* is in light tones. The darker tones in the three images emphasize the water in the rice fields

for its economic cost as well as for the complexity of the system and its scarce flexibility and operability.

In the SAR technique, the sensor is fixed on satellite platforms, carried by aircraft or on the ground. Geometric resolution of the obtained images ranges in 10–25 m for current operating satellites and 1–3 m for more advanced sensors carried by aircrafts and for the most recent TerraSAR-X and COSMO/SkyMed satellite systems. In

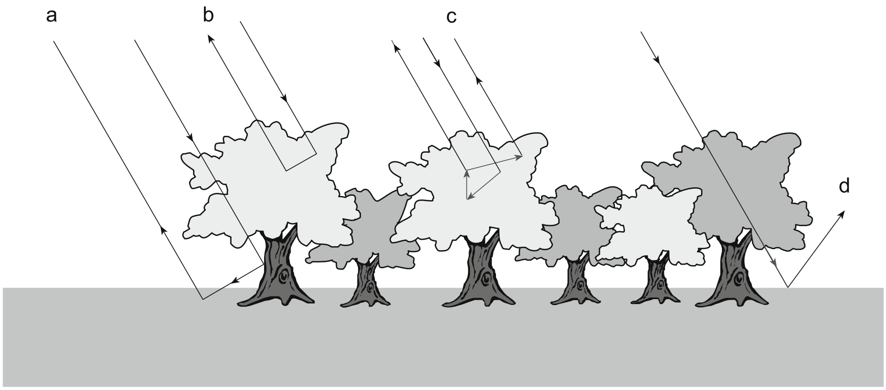


Fig. 4.30 Different possible radar backscattering signals in a woody area: (**a**) angular, (**b**) leaf surface, (**c**) biomass volume and (**d**) penetration in the biomass

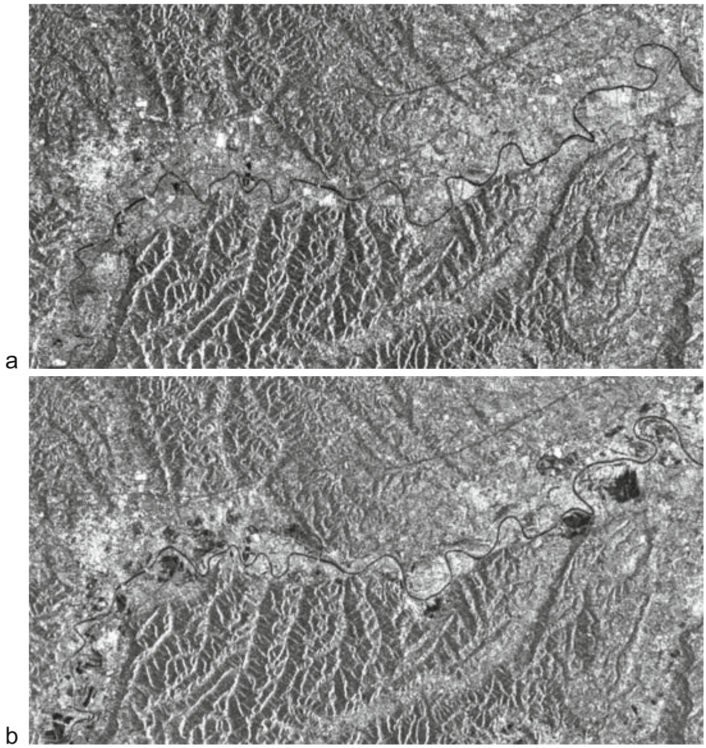


Fig. 4.31 Comparison between two SAR ERS-1 images of October 4th and November 9th before and after an alluvial event in Piedmont, Italy. In *black*, the flooded areas along the Tanaro river course

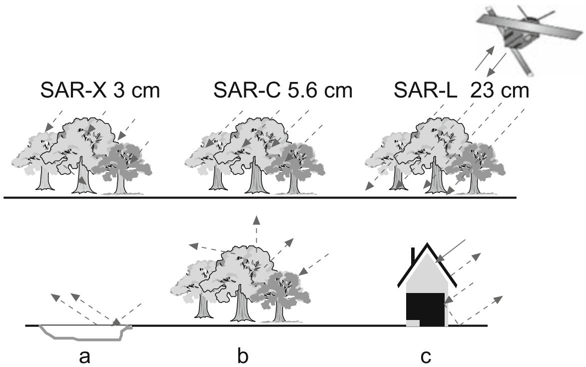
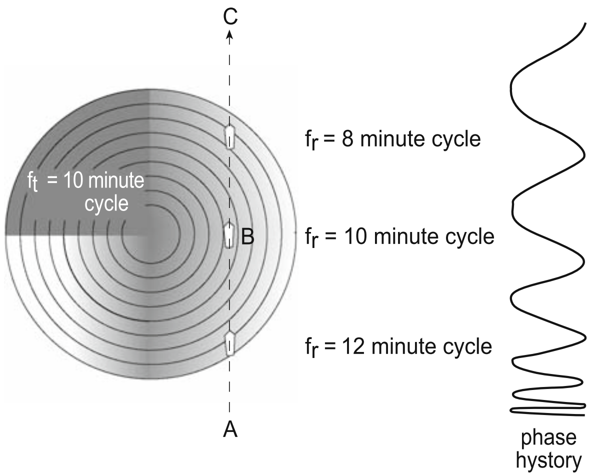


Fig. 4.32 The depth of penetration of the radar pulse is a function of the wavelength. The reflected energy (backscattering) is low in the presence of the water body, intermediate for vegetation, high in the urban area

ground sensors, geometric resolution depends on the distance from the detected area to the sensor. Centimetric geometric resolutions can be obtained by sensors that are tens of metres from the acquisition point and metrical geometric resolution by sensors that are kilometres from the acquisition point.

Fig. 4.33 Illustration of the Doppler effect: a boat moves from A to C in uniform rectilinear motion in a wavy environment with constant wave frequency f_t. On board the boat a different wave motion f_r is registered, at a frequency that differs from f_t except in position B. The absolute difference between f_t and the recorded frequencies f_r is defined as Doppler frequency



The system that brought a revolution in technology is the SAR: radar with a short antenna, hence easy to carry on board a satellite, able to act like a longer antenna and thereby to obtain very good resolution of the ground image. This result can be obtained by using the *Doppler Effect* (Fig. 4.33) while the satellite moves and by appropriate techniques of data processing (Ulaby et al., 1981, Prati, 2000).

Studies for environment observation and monitoring through spatial remote sensing techniques in radar band with SAR instruments were carried out in the early 1970s: the Seasat-A satellite, operating in orbit in 1978 for 105 days, for the purpose of monitoring the ocean surface and to provide radar images of the Earth's surface.

Based on this first positive experience, many satellites with radar image acquisition systems went on to be built.

The characteristics of the instruments for acquisition in radar band from satellites are described in Chapter 6.

4.4.2.1 Typology of Radar Sensors

Radar sensors are distinguished based on the typology of data they can produce:

- *generating images (imaging)*: i.e. *Real Aperture Radar* (RAR), *Side Looking Airborne Radar* (SLAR) and *Synthetic Aperture Radar* (SAR);
- *non-generating images (non-imaging)*: sensors collecting measurements only; i.e. the radar altimeter, the microwave scatterometer.

In imaging radars, the radar impulse is transmitted out of the vertical (nadir) by an antenna fixed beside the platform. It is not transmitted towards the nadir to avoid ambiguity about the position of the targets symmetrically located on the left and on the right of the platform. These targets would produce echoes that simultane-

ously reach the radar antenna and could not be distinguished. The echoes of the signal are processed in order to generate an image where each pixel is clearer if the corresponding area is characterized by a high backscattering signal, and the amplitude of the radiation emitted by the antenna is directly proportional to the impulse wavelength and inversely proportional to the antenna length.

The RAR cannot generate high-resolution images, as the aerial or satellite platforms cannot physically support very long antennae.

In the case of SAR, resolution is increased making a known size antenna move along a definite distance: it will record the back signal for that distance, acting like a longer antenna, but with the advantage of avoiding the transportation of a larger antenna. SAR techniques use signal processing methods obtaining resolutions independent of the distance and use a small-size antenna, through the *Doppler Effect*.

The process consists of storing and comparing the time that the back signals take to arrive from the detected surface, while the sensor moves along the flight line. The numerous sent signals have constant frequency while the back signals have a variable frequency, as antenna and detected object are in relative motion in respect to one another (Fig. 4.33). Processing the entire back signals, the effect of an aperture antenna is simulated as a much larger size than really exists.

Using this technique, ERS-1 and ERS-2 with an antenna 10 m long carry out detection corresponding to a 5 km real antenna.

Non-imaging radar instruments are

- *radar altimeter*: it examines the backscattered signals; nadir pulses with incidence angle perpendicular to the Earth's surface are emitted. Ocean and marine waves' height can be monitored and topographic survey performed. In recent years, the radar altimeter has been experimentally used in monitoring the increase of sea level as a consequence of global warming. In particular, data collected by two radar altimeters on board the satellite TOPEX/Poseidon (launched in 1992) were analysed: they recorded increasing of the sea level of 2 mm/year (Douglas, 1997, Bindoff NL et al., 2007). This result is comparable with the mareographs measures located in areas not subject to errors due to terrestrial crust lifting (bradyseism);
- *microwave scatterometer*: measures the scattering and the reflecting capacity of a surface detecting an area through two or more identical instruments and in different directions. They provide information about wind direction and speed blowing over the sea surface, based on water roughness;
- *microwave radiometer*: also measures low brilliance temperatures; the radar signal penetration depth increases as the wavelength increases and as a function of the detected surface. The resulting intensity depends on the quantity of energy backscattered by the elements (*scatterers*).

4.4.2.2 Deformation of Radar Images

Besides the well-known advantages, radar images are subject to some distortions, mainly *geometrical distortions* and textural distortions (*speckle*).

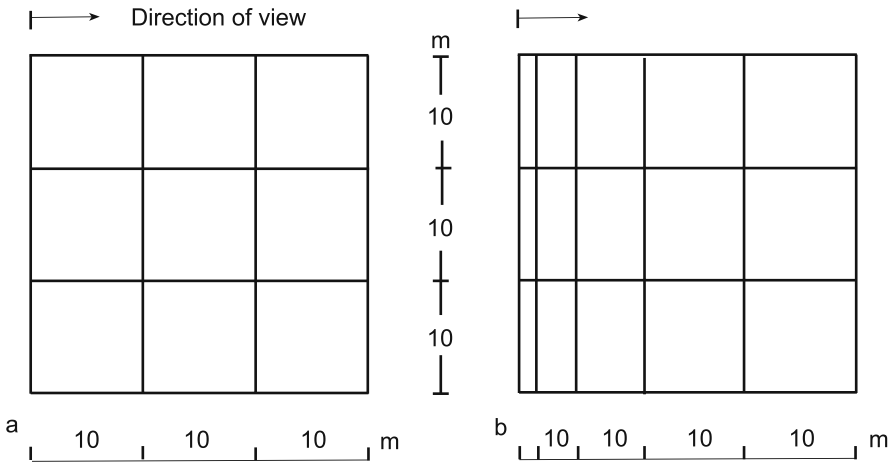


Fig. 4.34 Geometric comparisons of the optical and radar systems: (**a**) geometry of an optical vertical acquisition (nadir) and (**b**) geometry of a radar acquisition (slant range)

- *Geometrical distortions*: the two spatial dimensions of the SAR image are related to the distance of the objects from the sensor (*slant range*) and to the platform position along the motion direction (*azimuth*). Due to this representation, SAR images have the same geometrical deformations as images acquired by an optical system observing the terrestrial surface with a complementary angle of view. Ground objects on a slope comparable with the angle of sending signals by radar antenna, i.e. parallel to the antenna itself, are all at the same distance from the radar. These objects are represented in the same resolution cell without any possibility of discriminating them, even if very far from one another and if they are included in the plane with this slope. The same kind of geometric distortion is obtained in an optical system observing the plane from a complementary angle of view, like in a postcard seen from the side (Fig. 4.34). These zones are called *foreshortening zones*. In the optical case, this disposition gives a better geometric resolution. The contrary happens for grounds with opposite slope, where the SAR gives a better resolution and the optical system *flattens* the objects in one point of the image. This is due to the time of radar operativity.
- *Speckle*: in radar images, the concept of texture is essential: texture is the geometric distribution of superficial macro-roughness (Box 4.7). It is related to the dimension and average spatial organization of the elementary cells, which are the smallest homogeneous elements with the same radiometry. There are three levels of texture analysis: micro, meso and macro scale.

Speckle, characterizing all radar images with a grained effect or granularity (salt–pepper effect), is the statistical fluctuation associated with the pixels (see Fig. 4.28). In fact, within the zone containing many scatterers, a phenomenon of constructive or destructive interference occurs casually, originating light or dark tones in the image.

In image visualization, the speckle is a source of disturbance, but the quality can be improved by applying digital filters. In other applications, like SAR interferometry,

Box 4.7 The three components of the texture in radar images

Scale	Typology	Dimension	Characteristics
Micro	Speckle	Ct = Cr	- Always present in the imagery - Resolution independent - Random distribution
Meso	Texture	Ct > Cr	- Scene dependent - Resolution dependent - Concept of repetitiveness
Macro	Structure	Ct > > Cr	- Scene dependent - Resolution dependent - Repetitiveness not necessary

Ct: textural cell, Cr: resolution cell

speckle is useful and must be treated in a different way than what is done to improve the quality of visualized images.

4.4.3 Radar Techniques

Although radar spatial missions were born with the purpose to use the advantages of active sensors all-time operative capacities using intensity images, the main applications are based on techniques that use the phase differences between subsequent pulses.

Some possibilities in the use of radar images are (Fig. 4.35):

- *synthetic aperture radar images* (SAR), acquired in 100 km × 100 km scenes with a resolution of 5 m × 20 m (ERS-like systems), giving a very precise measure of the distance from the satellite and the observed points equal to about 1 every 100 m^2;
- *satellite SAR interferometry*, a technique using the observations of a same scene from two parallel and not coinciding orbits generating a Digital Surface Model (DSM) as distance's difference between the two satellites' positions and each ground point, with a vertical precision of about 10 m;
- *satellite SAR differential interferometry*, a technique using the observations of a same scene from the same orbit covered by the satellite in subsequent times, generating a map of relative movements of the observed points as difference between the distances from the satellite, with a vertical precision of about 1 mm and horizontal precision of about 1 cm.

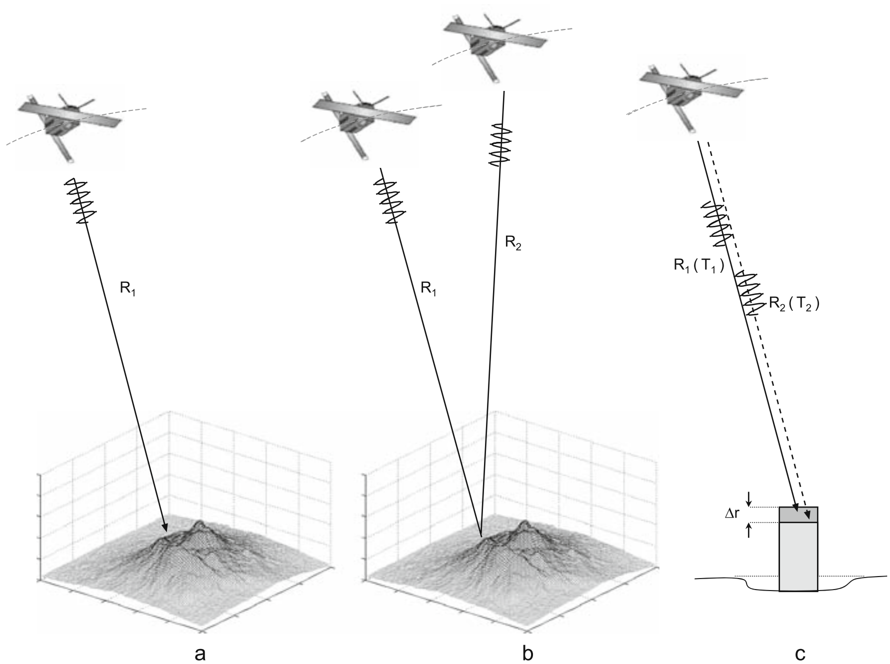


Fig. 4.35 Techniques of utilization of radar images: (**a**) single acquisition, (**b**) tandem (2 SAR on board 2 satellites) acquisition on different orbits, possibly at the same time, for interferometric application, (**c**) acquisition from the same orbit (1 SAR 1 satellite) at different times for differential interferometry

4.4.3.1 Radar Interferometry

The acquisition of a scene from different angles through radar instruments (*Synthetic Aperture Radar*, SAR) led to the development of the Differential SAR Interferometry Technique, called D-In-SAR. This technique retrieves the Earth's surface deformations much smaller than 1 cm by processing the phase information of the backscattered signal (Prati et al., 2000).

The differential technique consists of comparing two radar images acquired by two antennae during two orbits separated by a known distance, i.e. the baseline. Two microwave electromagnetic signals are emitted, characterized by a phase difference that is a function of the distance between the two antennae, corresponding to the base of the interferometric acquisition (Fig. 4.36). The difference between the phase values for each pixel corresponding to the same ground portion in the two acquisitions produces an interferogram, i.e. an image containing only information about the phase differences at different distances of the pixels from the SAR antenna.

The difference between the distance of each resolution cell, or pixel, and the two relative positions with respect to the radar antenna estimate and point out move-

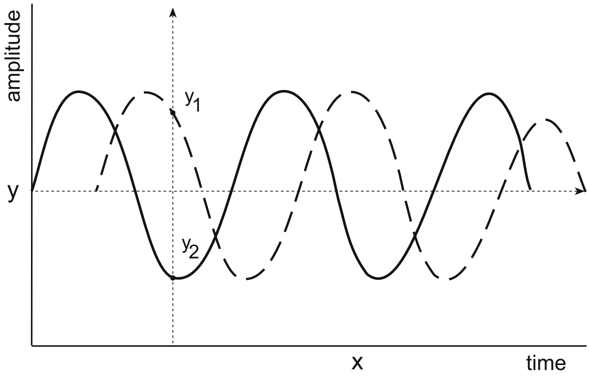


Fig. 4.36 The principle of the radar interferometry is based on the phase difference: two sinusoidal waves with same wavelength differ between y_1 and y_2 for a specific value x attributed to the phase difference between the two waves

ments or modifications of dislocated elements on the ground and/or about the ground itself.

One of SAR characteristics is its capability to retrieve *coherent* images, i.e. containing information about:

- *intensity*: radiometric width, expressing the object reflectivity at the observation frequency, in relation to the characteristics of the backscattering;
- *phase*: depends on the sensor–object distance, or length of the path from the radar to the target, on the backscatterer (the object) dielectric characteristics and, less, on the dielectric characteristics of the medium (air) in which the propagation occurs.

The use of the phase information contained in SAR image, sometimes combined with the radiometric information, opens up other possible advanced applications of radar observations, among which SAR interferometry.

Whereas the geometric resolution is of the order of metres, the phase measure, hypothesizing it only depends on the position, allows precision of the order of a fraction of the used wavelength, sometimes reaching a tenth of a millimetre.

The radar signal can be represented by a complex number α having amplitude A and phase ϕ according to the relation:

$$\alpha = A \cdot ej\varphi = \tag{4.14}$$

where e = 2,71828182..., Euler's number; j: imaginary number j=–$1^{0.5}$;j is a complex number whose squared value is a negative real number (not greater than zero).

Each pixel of a SAR image is associated with a complex number resulting from the combination of the backscattering of all the objects in a ground resolution cell and the rotation phase related to the path.

In particular, each pixel phase results from the sum of two elements:

- the first ϕ_σ depending on the backscatterers;
- the second ϕ_ρ given by:

$$\varphi_p = 4\pi\rho/\lambda \tag{4.15}$$

where ρ : platform–resolution cell distance; λ : radar wavelength (equal to the propagation velocity divided by the radar frequency).

Since the generally used wavelengths are equal to a few centimetres and the sensor–resolution cell distance is 800 km for ERS-1 and ERS-2, the second term of phase contains millions of giro angles. Moreover, the phase term related to the backscatterers is casual. As a consequence, the phase of a single SAR image is unusable.

On the contrary, considering the phase difference between two SAR images acquired from slightly different angles of view, the phase term due to the backscatterers can be disregarded if the angle difference is very small. In this case, the residual phase term is given by:

$$\varphi = 4\pi\,\Delta\rho/\lambda \tag{4.16}$$

where $\Delta\rho$: difference of the path from the sensors to the same ground resolution cell.

The phase φ still contains a very high number of giro angles (thus it is known for less than a high multiple of 2π); nevertheless moving from a resolution cell to the next one (a few metres distance), the variation of φ is small enough not to show ambiguity of 2π. The phase φ is said to be *interferometric phase* and the information about variations of $\Delta\rho$ (measured in fractions of wavelength) among SAR image pixels is related to it.

SAR interferometry hence is based on the phase as a physical property of the electromagnetic wave.

Even if the technique is applicable to pairs of images acquired with slightly different angles of view, the simplest configuration, but not optimal if more images are available, to measure ground deformations prescribes the acquisition in different times and exactly from the same position (null baseline). In these conditions, considering the phase difference between corresponding pixels in the two images and assuming that the dielectric contribution is constant and therefore becomes null, the topographic contribution to the phase variation is null, while the residual value can be correlated to a distance variation. This way it is possible to measure movements of the order of the used radiation wavelengths, usually from the centimetre to the millimetre, with accuracy equal to a small fraction of that wavelength. It should be noted that the typical ambiguity in phase measurement is only partially solved as the relative movement corresponding to each pixel does not have to exceed the half wavelength – factor 2 relates to the path sensor-object being covered both on the way to and the way back.

Integrating data obtained by the interferogram with the altitude of the two antennae the Digital Surface Model (DSM) can be obtained.

The phase of the data allows the SAR to generate the DSM with a precision of few metres and to be a unique instrument for providing measures of crust deformation of wide areas, even hundreds square kilometres, with centimetric precision and high density, equal to a measurement of every few tens of metres on the ground.

In 1995 ERS-1 and ERS-2 were placed on the same orbit in order to operate over the same zone at a one-day distance. With this disposition (*tandem*) of the satellites, data about the whole Earth's surface were acquired, which now are a database generating a DSM of large areas of the terrestrial surface. Once the position of the two satellites is known, $\Delta\rho$ measure can be used to retrieve the relative altitude of the image pixels, and then generate a numerical model of elevation (DSM). Or, knowing the DSM, it is possible to retrieve, from Δr, millimetric deformations of the terrestrial surface that occurred between two subsequent observations.

$\Delta\rho$ measured precision is related to the phase noise in the SAR images. If SAR images used to calculate the interferometric phase are acquired simultaneously (therefore the phase contribution of the backscatterers becomes null), the phase noise is generally lower than 30° and $\Delta\rho$ measured precision is better than $\lambda/20$ (a few millimetres). DSM is retrieved from $\Delta\rho$ measure and the parallax, sensibility strongly decreasing, with a few millimetres errors on Δr and errors of some metres on DSM. If SAR images are acquired within a temporal gap (this is the case of both SEASAT and ERS-1/ERS-2 satellite interferometry), the phase noise mainly depends on ground backscatterer changes, and $\Delta\rho$ measured precision has a strong spatial variability.

Whereas this variability is an inconvenience for some applications such as DSM generation or measurement of crust deformations, it can be an advantage for image classification and for extracting some geophysical parameters (*coherence* images generation for these kinds of applications).

Coherence Images

If ground backscatters change within two subsequent SAR observations, the interferometric phase is affected by a casual noise. The entity of this noise is evaluated through coherence images i.e. by the assessment of the SAR images local cross-correlation coefficient. In theory, each pixel of the images is associated as a different coherence value, but in practice, since only two images are available for coherence assessment, therefore the signal is supposed to be stationary within a few pixels area. Coherence image resolution is then lower than the resolution of the starting image. Coherence ranges from 0 (completely different backscatterers in the two original images, for example, in the case of the sea or a biomass) to 1 (same backscatterers in the two images, for example, in the case of two exposed rocks or infrastructures).

Sources of coherence errors are all phenomena inducing abrupt or chaotic phase variations in the observed scene and in particular in the time span within the two acquisitions.

Effects determining loss of coherence are sudden variations of the dielectric constant, i.e. in the case of hard atmospheric precipitations, or sensible chaotic movements such as over densely vegetated areas or water bodies. Considering a densely vegetated area, where the entity of the movements is to be related to the used radiation wavelength, many portions of the tree canopy, also in very moderate wind conditions, produce relevant variations in the leaves' position with chaotic spatial movements such as to destroy coherence.

Fig. 4.37 Radar reflectivity map of Campi Flegrei, Napoli, Italy. Radarsat data, Canadian Space Agency

In the case of dielectric effects, it is possible that a spatially coherent and gradual variation in soil humidity appears coherent and thus can be confused with a movement, as virtually non-discernable.

An example of coherence image is shown in Fig. 4.37. The image shows the northeast part of Sicily, and the coherence is represented with a grey scale, from black (coherence equal to 0) to white (coherence equal to 1). The data that were used were acquired by the satellites ERS-1 and ERS-2 in a one-day span in September 1995. The sea is completely incoherent, whereas on the island there are many coherence levels such as low in the north part indicating more vegetated, and high in the Etna area in correspondence with lava flows, which was identified with a high certainty. Coherence varies also in function of the climatic condition, in vegetated areas generally being higher in dry periods.

Coherence of a Short Period

Besides the coherence treated above, concerning changes occurring within two acquisitions, there is also *short-term coherence*, referring to any change occurring in the observed scene during the single SAR image acquisition. SAR focalization algorithms, passing from raw acquired data to the final image, assume that the scene is fixed; targets having relevant movement within the time spent to acquire the set of raw data are not properly focalized (Prati et al., 1998). The error in focalization depends on the entity and type of the object movement in relation to the total time of measurement and can lead to effects ranging between a slight defocalization (i.e. in

the case of intense objects, part of the energy is distributed also over closest pixels) and the total disappearance of the object, corresponding to a chaotic distribution of the backscattered energy by the object over the whole image. Hence the first effect on SAR image is that moving targets are not sensed in the correct position and appear less intense.

With regard to the interferometric product, a second effect, which may be even more important is that some pixels receive decorrelated energy, i.e. energy not having useable phase information, coming from other areas of the observed scene and representing only noise for them. To evaluate the entity of the effect, the typical time during which the object movement called decorrelation time is realized and must be evaluated in relation to the time of measurement of the used radar system. In particular, it is to be assessed if decorrelation time is lower than the time needed for a single radar acquisition or only for the time for total acquisition of single image measurements. The two cases differ for the way in which decorrelated energy is distributed in the final image.

In the case of satellite sensors, the effect can be disregarded, as the measure time is about 1 s. For other sensors, i.e. ground sensors, this time is much longer and cannot be disregarded.

4.4.3.2 Digital Surface Model (DSM) Generation

From the interferometric phase, it is possible to retrieve a map of relative altitude of all the pixels belonging to a scene in a definite zone with certain morphologic characteristics. Two operations are required to obtain this result:

- de-rolling the interferometric phase (phase unwrapping);
- DSM georeferencing.

The first operation is needed since the interferometric phase has 2π gaps not relating to a real difference in pixel altitude but depending on the phase representation which is known for less than 2π multiples.

Using more images, the percentage of zones having high coherence in one of the two situations is increased and, consequently, DSM of wider areas can be obtained. Once an elevation map in SAR coordinates is obtained, this has to be inserted in a conventional reference system (usually UTM) by georeferencing operation. Figure 4.38 shows an example of DSM in the Etna area generated from seven pairs of SAR images.

Due to geometric deformations of SAR images, foreshortening zones are strongly interpolated and not very reliable. In order to obviate this inconvenience, elevation maps obtained with SAR images pairs acquired during both ascending (from south to north) and descending (from north to south) paths of ERS-1 and ERS-2 satellites were combined. Geometric deformations in the two cases are almost complementary (in ascending paths the antenna points eastward, in descending paths westward); almost the whole surface of a mountainous area like Etna area is represented with good detail.

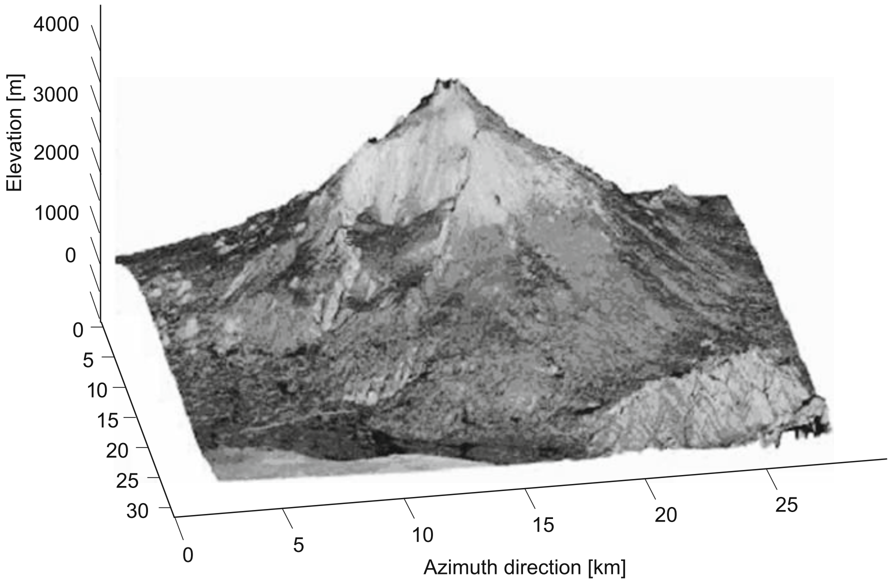


Fig. 4.38 Etna Vulcano Digital Surface Model derived from seven interferometric pairs of ERS-1 and ERS-2 SAR images

A non-negligible problem in DSM generation through non-simultaneous SAR images (like in the case of ERS-1 and ERS-2) is water vapour content variations in the troposphere within two subsequent observations and/or between two zones in the same observation. These phenomena cause local variations in the system wavelength and consequently topographic artefacts; using more interferometric images helps to reduce these effects. For this reason, a technique based on SAR interferogram multi-resolution processing (bi-dimensional wavelet filtering) has been developed for assessing the atmospheric artefacts on single interferometric couples. Based on these evaluations, it is possible to find the best linear combination of the single interferometric pairs to reduce atmospheric effects to the minimum.

4.4.3.3 Applications of Differential Interferometry

One of the main applications of differential interferometry is in the field of monitoring environmental risk and phenomena inducing ground movements. In relation to the used platform, both phenomena that can be studied and the spatial–temporal scale at which they are observed are different.

In the case of satellite platforms, there are a lot of positive examples showing that this technique is unique for measuring wide areas of subsidence due to different causes, like earthquakes, mining, volcanic activity. Through images provided by current operative satellites, precision of the order of a centimetre is reached, and the typical resolution is of the order of 15–20 m (Plate 4.10).

More recently a technique called Permanent Scatterers (PS) has been proposed and successfully applied (Rocca & Prati, 1998, Ferretti et al., 2000): it applies very well to the study of long-term movements especially in urban areas and in the presence of exposed rocks and non-vegetated terrain. Resolution is the same as satellite images, whereas precision is sub-centimetric (see Plate 4.7). Through the technique of Permanent Scatterers, unprecedented instruments for measuring millimetric deformations of terrestrial surface were developed and patented in 1999 by the Department of Electronic and Information of Polytechnic of Milan.

If topography is known, its contribution to the interferometric phase can be eliminated. Interferometric phase residue can be related to small relative movements of the terrestrial surface in the satellite direction (Fig. 4.39). In the case of satellites ERS-1 and ERS-2, for example, a relative movement of 2.8 cm, that is half the system wavelength, causes an interferometric phase variation of 2π. With high coherence in the zone of interest, movements of a few millimetres can be measured (Plates 4.8 and 4.9).

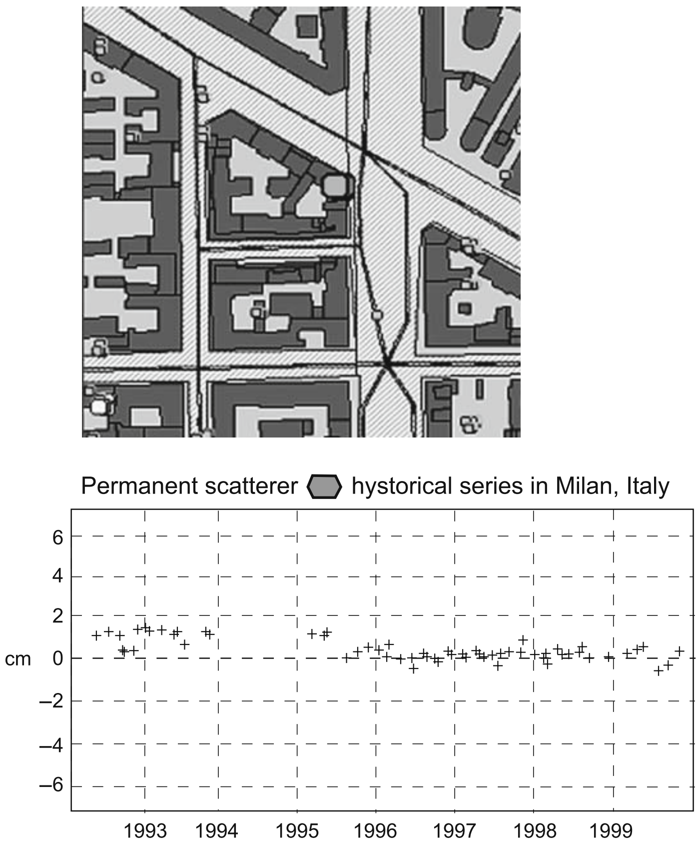


Fig. 4.39 Temporal (1992–1999) trend of the reference point on the map. Every cross in the plot corresponds to one radar image with application of the technique of the permanent scatterers *(DEI, Politecnico of Milan, Italy)*

The main limitations to fully operative use of satellite images depend on the fact that mission characteristics, both in terms of orbital parameters (such as acquisition frequency on a certain area) and instrumental parameters (such as incident angle and observation frequency) are fixed and thus not adaptable or perfectible for a definite application. Produced images can even result in being inadequate, especially in terms of geometric resolution, or unusable if the phenomenon of interest occurs on steep slopes. On the other hand, satellite data offer the relative advantage of covering the whole of the Earth's surface providing 100 km $\times$ 100 km typical extension single acquisitions. Moreover in the case of satellites such as ERS, a large archive of images since 1992 is available.

There are many cases in which interferometric techniques with satellite data were successfully applied to studying landslides (Fruneau et al., 1996; Kimura & Yamaguchi, 2000). Although in many cases satellite data do not allow an accurate study of a definite phenomenon, they can be used for quick alert purposes as they cover wide and remote areas.

The use of sensors carried by aircraft offers more flexibility in terms of both instrumental parameters and acquisition geometry. The main technical problem depends on low stability of the aerial platform, which is reflected in instrumentation complexity and in the complex data correction algorithms required. Whereas for collecting topographic data, the aerial platform gives very good results, but in regard to differential interferometry it has not succeeded in improving the centimetric precision already obtainable by satellite data.

4.4.3.4 In-Field Differential Interferometry

The use of ground sensors has many advantages in landslide monitoring, as the characteristics of these phenomena are very variable (Tarchi et al., 2000). These sensors offer flexibility, enabling adaptability in any situation, in terms of:

- observation frequency;
- acquisition geometry;
- polarization;
- multi-frequency of acquisition; and
- the short time in which they can be operative.

Tri-dimensional geometries are also possible in carrying out tomographies of the examined area.

In the case of ground sensors for differential interferometry applications, it is technically easier to realize the condition in which the acquisition is repeated exactly from the same position. Combining this aspect with the possibility to reach millimetric observation frequency, the measure precision of movements can be submillimetric. The main limitations are the maximum width of the covered area, some square kilometres, and the fact that applicability depends on the possibility of finding a good observation point, especially on flat areas.

This technique allows applications in monitoring movements of big infrastructures such as dams, bridges or historical–artistic buildings. In the latter field, radar interferometry is a good integration to traditional methods, with the possibility of offering data about the area together with the advantage of acting remotely and without needing intervention on the buildings.

4.5 Summary

Remote Sensing (RS) is a powerful tool for environmental analysis, relying mainly on the advantages given by a synoptic vision, a repetition cycle enabling time-changing studies and multispectral acquisition capabilities. The origins of this discipline come from photography, optical, and spectrometry studies being deeply rooted in the domains of the electromagnetic spectrum and aeronautical sciences.

The electromagnetic spectrum of radiation wavelengths can be divided into different regions having different properties and can be investigated through different remote sensing tools and techniques. They are ordered by increasing wavelength: cosmic rays, gamma (γ), X rays, ultraviolet, *visible light* (ranging within the interval of human eye spectral sensibility), infrared, microwaves and radio waves.

Up to between 2.5 and 5 μm wavelength, the electromagnetic radiation from the Earth's surface is coming from solar energy reflected by the Earth, whereas for wider wavelengths the main contribution is coming directly from energy emitted by our planet into space, most of which is in the form of thermal energy.

The electromagnetic radiation coming from the Sun and passing through Earth's atmosphere undergoes the phenomenon of attenuation due to both the distance travelled and the absorption and scattering processes occurring in atmospheric transit and depends upon the wavelength, and that fact has to be taken into account when dealing with remotely sensed data.

The basic laws of Remote Sensing are those of electromagnetic radiation properties, and in particular the following: the *Kirchhoff's radiation law* which regulates the relationship among the coefficients of reflection, transmission, absorption and emission; *Planck's radiation law* which defines the behaviour of the energy emitted by a surface as a function of the wavelength, and the temperature; the *Stefan–Boltzmann's radiation law* which furnishes the total quantity of energy emitted by a surface calculated on the whole electromagnetic spectrum, for any temperature; and *Wien's displacement law* which points out the wavelength value in correspondence with the maximum electromagnetic emission at a defined temperature.

In particular, within the visible light range, one of the most important properties investigated through remote sensing has been the colour of objects and surfaces: the colour is in fact the visual perceptual property fundamental to describing the external world and physically originated by coloured lights and/or coloured pigments. It interacts in the human eye, our basic remote sensing instrument, with the spectral sensitivities of light receptors. Numerous efforts have therefore been put into place to classify and describe colours.

The concept of radiometry and radiometric magnitudes are basic, especially in the case of optical and passive remote sensing measurements. The *radiance* is the

most important of radiometric units, as it describes what is measured by the sensors used in remote sensing in the real world.

The optical-passive sensors on board satellite and aerial platforms or field instrumentation are sensitive to only one radiometric unit: the *radiance*. The signal recorded by the sensor can be represented in graphical terms as a reflectivity function of the wavelength: this is the *spectral response* of an object, which is a key instrument for quantitative analysis of satellite images.

Other than optical sensors, an important field for remote sensing is radar with active sensors that exploit microwave radiation for applications ranging from structural geology and hydrogeology surveys, to research about the sea surface, to volcanic, alluvial and glacial morphology, etc.

Both optical and radar instruments show different powerful analysis capabilities that are different but sometimes overlapping fields of application making them an even more powerful analysis tool when used in conjunction and in complementary ways.

An innovative application for radar data is the use of SAR ((*Synthetic Aperture Radar*) in interferometric and permanent scatterers techniques. These techniques retrieve Earth surface relief and deformations with high precision and reliability, derived from Digital Surface Models (DSM) or calculating displacement velocity.

Further Reading

Optical Passive

Campbell J.B., 1996, Introduction to Remote Sensing, 2nd ed. The Guilford Press, London.

Cracknell A.P., Hayes L., 2007, Introduction to Remote Sensing, 2nd ed. CRC Press, Taylor & Francis Group, Boca Raton, London, New York, Leiden, p. 352, ISBN: 978-0-8493-9255-9.

Landgrebe D.A., 2003, Signal Theory Methods in Multispectral Remote Sensing. Wiley, p. 528, ISBN: 978-0-471-42028-6.

Richards J.A., Xiuping J., 1998, Remote Sensing Digital Image Analysis: An Introduction, 3rd revisited and enlarged edition, Springer.

Radar

Ghiglia D., Pritt M., 1998, Two-dimensional phase unwrapping. Theory, Algorithms and Software. John Wiley & Sons, New York.

Ketelaar V.B.H., 2009, Satellite Radar Interferometry, Subsidence Monitoring Techniques, Springer, Berlin.

Ulaby F.T., Moore R.M., Fung A.F., 1981, Microwave remote sensing: active and passive, Vol. I. Microwave Remote Sensing, Fundamentals and Radiometry. Addison-Wesley Publishing Company, Reading, Massachussetts, USA.

Ulaby F.T., Moore R.M., Fung A.F., 1982, Microwave remote sensing: active and passive, Vol. II. Radar Remote Sensing and Surface Scattering and Emission Theory, Addison-Wesley Publishing Company, Reading, Massachussetts, USA.

Ulaby F.T., Moore R.M., Fung A.F., 1986, Microwave remote sensing: active and passive, Vol. III. From Theory to Applications. Addison-Wesley Publishing Company, Reading, Massachussetts, USA.

Woodhouse I.H., 2006, Introduction to Microwave Remote Sensing. CRC Press, Taylor & Francis Group, Boca Raton, London, New York, Leiden, p. 400, ISBN: 978-0-415-27123-3.

Bibliography

Optical Passive

Bruzzi S., 1995, Special Feature: ENVISAT. EARSeL Newsletter, 4: 4–13.

Colwell R.N. (Ed.), 1983, Manual of Remote Sensing, 2nd ed. American Society for Photogrammetry and Remote Sensing, Falls Church, Virginia.

Gomarasca M.A., 2000, Introduzione a Telerilevamento e GIS per la Gestione delle Risorse agricole e Ambientali, Ed, 2nd Ed. AIT, p. 250, 32 Tavole a colori.

Reeves R.G., Anson A., Landen D. (Eds.), 1975, Manual of Remote Sensing, 1st ed. American Society for Photogrammetry and Remote Sensing, Falls Church, Virginia.

Richards J.A., 1986, Remote Sensing Digital Image Analysis: an Introduction, Springer-Verlag, Berlin.

Swain P.H., Davis S.M., 1978, Remote Sensing: The Quantitative Approach. Mc Graw-Hill Book Company, Paris, New York.

Teillet P.M., 1997, A status overview of earth observation calibration/validation for terrestrial applications. Canadian Journal of Remote Sensing, 23(4): 291–298.

Thome K.J., Markham B., Barker J., Slater P.N., Biggar, 1997, Radiometric calibration of Landsat. Photogrammetric Engineering and Remote Sensing, 63(7): 853–858.

Radar

Atzeni C., Canuti P., Gasagli N., Leva D., Luzi G., Moretti S., Pieraccini M., Sieber A., Tarchi D., 200 l, Monitoring unstable cultural heritage sites with radar interferometry. In: UNESCO/IGCP Symposium on Landslide Risk Mitigation and Protection of Cultural and Natural Heritage, K. Sassa (Ed.), Tokyo, Japan, 15–19 January, pp. 257–264.

Bindoff N.L. et al., 2007, “Observations: Oceanic Climate Change and Sea Level”, Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press Cambridge, UK.

Brofferio S., Cafforio C., Rocca F., 1984, Communication Experiments in the SIRIO Program. Alta Frequenza, Numero speciale sull’esperimento SIRIO.

Douglas B.C., 1997, Global sea rise: a redetermination. Surveys in Geophysics, 18: 279–292.

Cafforio C., Prati C., Rocca F., 1991, SAR data focusing using seismic migration techniques. IEEE Transactions on AES, 27–2: 194–207.

Canuti P., Casagli N., Leva D., Moretti S., Sieber A., Tarchi D., 2002, Some applications of ground-based radar interferometry to monitor slope movements. International Symposium on Landslide Risk Mitigation and Protection of Cultural and Natural Heritage. Kyoto, Japan, 21–25 January 2002, UNESCO/IGCP-425: pp. 357–374 (CNR GNDCI: 2506).

Colesanti C., Ferretti A., Novali F., Prati C., Rocca F., 2003, SAR monitoring of progressive and seasonal ground deformation using the permanent scatterers technique. IEEE Transactions on Geoscience and Remote Sensing, 41(7): 1685–1701.

Curlander J.C., McDonough R.N., 1991, Synthetic Aperture Radar: Systems and Signal Processing. John Wiley & Sons, New York, p. 672.

Delancourt C., 1996, Observation and modelling of the Saint- Etienne-de-Tinee landslide using SAR interferometry. Tectophysics, 265: 181–190.

Dixon T.H., Amelung F., Ferretti F., Novali F., Rocca F., Dokka R., Sella G., Sang-Wan Kim, Wdowinski S., Whitman D., 2006, Space geodesy: subsidence and flooding in New Orleans. Nature, 441, 587–588.

ESA, European Space Agency, 1995, Earth Observation Missions, ESRIN, Frascati.

Eurimage, 1996, ERS, JERS-1 and Resurs-01 Missions, Training Courses on ERS SAR and Other Complementary Spaceborne Sensors, ESA-ESRIN, Frascati.

Ferretti A., Massonet D., Monti Guarnieri A., Prati C., Rocca F., 2007, InSAR Principles: Guidelines for SAR Interferometry Processing and Interpretation, ESA TM-19.

Ferretti A., Novali F., Burgmann R., Hilley G., Prati C., 2004, InSAR permanent scatterers analysis reveals ups and downs in San Francisco Bay Area, EOS. Transactions on American Geophysical Union, 85(34): 317–324.

Ferretti A., Prati C., Rocca F., 2000, Non-linear subsidence rate estimation using permanent scatterers in differential SAR interferometry. IEEE Transactions on Geoscience and Remote Sensing, 38(5): 2202–2212.

Ferretti A., Prati C., Rocca F, 2001, Permanent scatterers in SAR interferometry. IEEE Transactions on Geoscience and Remote Sensing, 39(1): 8–20.

Fruneau B., Achache J., Delacourt C., 1996, Observation and modelling of the Saint-Etienne-de-Tinee landslide using SAR interferometry. Tectonophysics, 265(3–4): 181–190.

Fruneau B., Achache J., Tarchi D., Nesti G., Prati C., Rocca F., 1998, Proceedings of PIERS Workshop on Advances in Radar Methods, pp. 230–233.

Gatelli F., Monti Guarnieri A., Parizzi F., Pasquali P., Prati C., Rocca F., (1994), Use of the spectral shift in SAR interferometry: applications to ERS-1. IEEE Transactions on Geoscience and Remote Sensing,32(4): 855–865.

Hilley G., Burgmann R., Ferretti A., Novali F., Rocca F., 2004, Dynamics of slow-moving landslides from permanent scatterer analysis. Science, Elsevier B.V., Amsterdam 1952–1956.

Kimura H., Yamaguchi Y., 2000, Detection of landslide areas using radar interferometry. Photogram Engineering and Remote Sensing, 66(3): 337–344.

Nico G., Leva D., Antonello G., Tarchi D., 2003, Elevation information derived from ground based SAR data. SPIE, Vol. 4883 Posa Editor, Bellingham, WA, USA.

Pieraccini M., Tarchi D., Rudolf H, Leva D., Luzi G., Atzeni C., 2000, Interferometric radar for remote monitoring of building deformations. Electronics Letters, 36(6): 569–570.

Prati C., Rocca F., Monti Guarnieri A., Damonti E., 1990, Seismic migration for SAR focusing: Interferometrical applications. IEEE Transactions on Geoscience and Remote Sensing, 28(4): 627–640.

Prati C., Rocca F., 1994, Process for Generating Synthetic Aperture Radar Interferograms. U.S. Patent N.5, 332, 999, July 26.

Salvi S., Atzori S., Tolomei C., Allievi J., Ferretti A., Rocca F., Prati C., Stramondo S., Feuillet N., 2004, Inflation rate of the Colli Albani volcanic complex retrieved by the permanent scatterers SAR interferometry technique. Geophysical Research Letters, 31, L 12606: p. 4.

Tarchi D., Rudolf H., Pieraccini M., Atzeni C., 2000, Remote monitoring of buildings using a ground-based SAR: application to cultural heritage survey. International Journal of Remote Sensing, 21(18): 3545–3554.

WegmŸller U., Werner C.L., NŸesch D., Borgeaud M., 1997, Retrieval of vegetation parameters with SAR interferometry. IEEE Transaction on Geoscience and Remote Sensing, 35(1): 18–24.

Zebker H.A., Goedstein R.M., 1986, Topographic mapping from interferometric synthetic aperture radar observations. Journal of Geophysical Research, 91: 4993–4999.

Chapter 5
Elements of Informatics

Computer science is a subject in continuous evolution which determines increasingly quick and close technological revolutions: to follow these evolutions is difficult. The matching with the numerous specific professional activities requires continuous updating, also because what is innovative today can become obsolete tomorrow.

The specialist in geomatics is influenced by the market and by the quantity of information and products apparently useful but sometimes unnecessary or not compatible with the organization of his work. A basic ability to consciously take the appropriate choice when necessary or possible is of fundamental importance.

A processor can just be used without having a precise idea of its functioning or can be programmed for a specific application; the instrument can be underestimated and considered able to reply to any upcoming query. The programming phase in function of the operator's short-, medium- or long-term needs is strategic in order to preserve the initial choice when new instruments are needed to update system limits.

A definition of computer science has already been given in Chapter 1, *Geomatics*. The comprehension elements, the logic, the tools and potentialities at the basis of this science are described in the following paragraphs. The purpose is not to give an exhaustive treatment about this topic or its basic elements. The attention is focused on the history, machine architecture, programming languages, algorithms, operating systems, software engineering, networking, database design, databases and their management systems (DBMS) and artificial intelligence as they constitute one of the essential elements of *Geographical Information Systems* (GIS) described in Chapter 9. In relation to the importance that the GIS are acquiring more and more as the Web distribution is spreading (WebGIS), Web architectures and infrastructures are described. At last the hardware and software *structure* for the geomatics are indicated.

M.A. Gomarasca, *Basics of Geomatics*, DOI 10.1007/978-1-4020-9014-1_5,

5.1 Milestones in the History of Informatics

The history of computer science passes through many discoveries and developments which go back to calculations by the *abacus* carried out by the Chinese since 2000 BCE and used later by the Greeks and Romans (Box 5.1).

A precursor of the modern universal arithmetical computers was called Mark 1, which was able to sum two 23 digits numbers in 6 s: it was August 1, 1944. Mark 1 was presented at Harvard University and represented a milestone in the world of computer science. The producer of that huge machine, 15 m long and 5 m high, 5 tonne in weight, more than 760,000 elements spread through 78 computers connected among them by 800 km of cables and electric wires, was the American physician Howard Aiken (1899–1973). He realized the old dream of the English mathematician Charles Babbage (1792–1871).

Box 5.1 Reference year, scientist and scientific discovery in the field of Informatics

Year	Scientist	Discovery
1642 –1645	Blaise Pascal	Invented the second mechanical calculator, called alternatively the Pascalina or the Arithmetique, a decimal machine that could add and subtract, the first being that of Wilhelm Schickard in 1623
1671	Gottfried W. Leibniz	Binary system: based on the symbol 0 and 1 and intuition on the possible use on computation machines
1804	Joseph Jacquard	Inventor of a loom which produced repeated weave patterns in cloth and carpets, programmed by a series of punched cards Later developed a machine where the punched cards were joined to form an endless loop that represented the program for the repeating pattern used for cloth and carpet designs
1822	Charles Babbage	Difference Engine No.1 was the first successful automatic calculator and remains one of the finest examples of precision engineering of the time. Babbage is sometimes referred to as *father of computing*
1836	Samuel F. B. Morse	Developer of *lightning wires* and *Morse code*, an electronic alphabet that could carry messages (electromagnetic telegraph) He realized that pulses of electrical current could convey information over wires

1854	George Boole	Boolean algebra, sometimes known as Boolean logic, is based on two-valued system, separating arguments into different classes which may then be processed according to the presence or the absence of a certain property, enabled any proposition – regardless of the number of individual items – to draw logical conclusions Boole's work has to be seen as a fundamental step in today's computer revolution
1890	Herman H. Hollerith	Inventor of a punch card tabulation machine system that revolutionized statistical computation. His system enabled the 1890 census to save $5 million and more than 2 years time The card had a dimension of a dollar
1890	Giuseppe Peano	*Sur une corbe qui remplit une aire plane:* first description of a curve that cover all the points in a square
1892–1900	Several authors	Use of the relay, constituted by a spool of conducing material and a metallic body, able to increase the memory capacity of the cards
1892–1908	Giuseppe Peano	Start and publication of the volume: *Formulaire de Mathematique*, 4200 formulas and theorems with written demonstration in symbolic form Fundamental for the vector computing theory: axioms of the Peano arithmetic
1915	Coadjutor of Guglielmo Marconi	Inventor of the valve, device created for the radio and able to accelerate the calculation processing Its use, jointly with diode, substituted the relay utilization, opening the way to the semiconductors
1924	IBM	From the fusion of the Hollerith Company with others minors was borne IBM, *International Business Machines Co,* starting the history of the electronic computing machines
1936	Alan Turing	Inventor of the universal machine, a programming calculator in the most modern interpretation
1944	Howard Aiken	First electro-mechanical calculator Mark 1
1945	John von Neumann	Theory report on the von Neumann machine, based on the fundamental elements: processing unit, central memory, peripherals and bus system. CPU slow-to-access storage area, like a hard drive; secondary fast-access memory (RAM). The machines stored instructions as binary values (creating the stored-program concept) and executed instructions sequentially – the processor fetched instructions one at a time and processed them The von Neumann publication ignores the contribution of J. Presper Eckert and John William Mauchly who worked on the concept

1952	John von Neumann	First realization of the von Neumann architecture distinguishing between primary (ROM) and secondary (RAM) memories and programming style using flux diagrams The term von Neumann architecture refers to a computer design model that uses a single storage structure to hold both instructions and data. The term von Neumann machine is used to describe such a computer, but that term has other meanings as well. The separation of storage from the processing unit is implicit in the von Neumann architecture The term *stored-program computer* is generally used to mean a computer of this design
1960	J.C.R. Licklider	Development of the theory of Internet at the Massachusetts Institute of Technology (MIT)
1972	Bill Gates and Paul Allen	They formed Traf-O-Data, 3 years later renamed Microsoft. In February 1975, the first two *hackers* presented the language Basic They had a vision of software as a force of crucial importance in the computer industry
1990		Start of Internet
1991		Start of World Wide Web

Innumerable discoveries followed, which, after the first generation of computers, led to the development of data storage processing, to the second generation with the invention of transistors, the development of languages to communicate with machines, up to the third generation that introduced integrated circuits, to arrive at the radical revolution in the 1970s and to the spreading of microcomputers and Personal Computers (*PCs*).

5.2 Architecture of the Computing Systems

The term computing systems is used to refer to objects at the extremes in respect to one another: from the smallest portable computer to the multi-user mainframe capable of managing huge amounts of data using very complex programs. These typologies of computing systems are at the boundaries of a continuous space which includes systems of increasing complexity; although very different, they have elements in common.

A computing system is a complex element made up of many interacting parts. To comprehend computing system architecture means to recognize each part of it, understand their functioning general principles and understand how the different components integrate.

The main distinction is between:

- *hardware*: the physical components of the system;
- *software*: the programs executed by the system.

Before going into details it is fundamental to define an algorithm, since computing systems are algorithm executors.

5.2.1 Algorithm

An algorithm can be informally described as a sequence of precisely defined steps which lead to the realization of a task. An algorithm necessarily has to be understandable to its executor. The algorithms are described by a program, i.e. a sequence of instructions written in an appropriate language understandable by the computer that executes them in a sequential and precise, quick and efficient way, although it is not able to interpret when the execution is evidently wrong. This is because the computer is not intelligent: it cannot complete or insert data not provided or instructions if not previously accounted and codified by the operator.

The computation of a Vegetation Index based on the ratio between the Digital Numbers of two different spectral bands belonging to the same pixel with x, y image coordinates is an example of an algorithm.

The fundamental characteristics of an algorithm are

- *exactness*: when it finds a correct solution to the problem assigned to it;
- *efficiency*: when it finds the solution to the problem using the minimum physical resources as possible and/or in the quickest way.

For example, a user can consult geographic maps to solve a problem, but the way he/she processes the information contained in the map does not always correspond to execute an algorithm. By knowing the path between two points, an algorithm can indicate and calculate the shorter way between them; but if along the path to reach a defined place there are conditions that cannot be seen on the map, the algorithm cannot substitute the map.

In approaching an example of constraint, the algorithm just executes the instruction without any critical approach: all the pixels of 8-bit infrared bands with Digital Numbers included between 120 and 150 belong to the class *vegetation*; although this sentence is not completely true, the algorithm gives a solution.

There are different programming languages, understandable by the computer, used to describe and codify the algorithms. They are distinguished in

- *high-level programming languages*: Fortran (*FORmula TRANslator*), particularly suitable to describe mathematical formulae processing, Cobol (*COmmon Business-Oriented Language*), suitable for management applications;
- *languages based on a study of the programming principles*: Pascal, used in teaching; C++, the evolution of the Object-Oriented (OO) C, whose tendency is to keep

a correspondence between the objects characterizing an application and their codification; Java, used for systems programming on computer networks;
- *non-conventional languages*: thought to avoid the influence of the programming style by the computation instrument characteristics using mathematical languages such as LIPS, based on the concept of function, and the Prolog, based on mathematical logic formalism.

Internal solutions in the machine have been realized, able to translate the program into the machine language. For example, the *editor*, for the construction of files containing texts, the *compiler*, for translating a source program into an object program written in a language executable by the processor, the *linker*, a link system between the different programs, or modules, coordinated among them, and a controller of the execution, the *debugger*, to eliminate any errors occurring in a program.

5.2.2 Computer Hardware

5.2.2.1 Digital Code

In a processor, data and instructions of a program are codified in binary form, i.e. in a finished sequence of 1 and 0. The smallest information unit that can be stored or processed by a computer, the *bit* (binary digit), corresponds to the status of a physical device that is interpreted as 1 or 0. Another important information unit is the *byte*, equal to 8 bit. A byte can be used to produce 2^8 different sequences of 1 and 0 (00000000, 00000001, 00000010, ..., 11111111).

A computer can process different kinds of data: natural, integer, real and fractionary numbers, texts, images, sounds, etc. All these have to be transformed into bit sequences to be processed. There are several conventions to codify data; for example to represent:

- *natural numbers*: 8 bit are sufficient, one byte, 2^8 (0–255) for their representation (0: 00000000, 1: 00000001, . . ., 8: 00001000, ..., 255: 11111111);
- *integer numbers*: include negative integer numbers, the zero and positive integer numbers, thus also their sign has to be reported. The first byte is used to attribute the sign. As convened, a 0 indicates a positive number while a 1 indicates a negative number. Thus 8 bit can represent the integers between –127 and +127, namely between $-(2^{(8-1)}-1)$ and $+(2^{(8-1)}-1)$;
- *real numbers*: rational numbers containing an integer part and a fractional part approximating the real number with arbitrary precision. If the notation in fixed point is used, a real number can be represented separately codifying its integer part and its fractional part. This case, for example, the number 8.88 would be represented in 2 base as follows:
  - first byte: 00001000, for the representation of the integer 8;
  - second byte: 01011000, for the representation of the fractionary part 0.88.

5.2.2.2 Von Neumann Architecture

The mathematician John Von Neumann (1903–1957) suggested the concept of *computer for general applicability*, an instrument for information processing able to manage different programs, performing its task and reaching different results, without modifying the device's physical structure.

In this tool, the processor also indicated as *Central Processing Unit* (CPU) has a central role as the exchange of information among the different units involved in the programs' execution.

CPU is a description of a class of logic machines that can execute computer programs.

The elaborator depends on digital technology, and Von Neumann defined general-purpose elaborator architecture according to the following guidelines:

- a programmable device requires a memory able to store both data and instructions;
- the memory has to be associated with an arithmetical logic unit, able to process data and instructions contained in the memory;
- the processing of data from the memory has to be mediated by a Control Unit;
- in order to interact with the user the device has to be adapted through instruments for communication with the outside world, i.e. acquisition and distribution (input/output) instruments.

According to these guidelines, the main schematic components constituting an elaborator can be identified:

- *processor*, or processing unit, or CPU, logic–arithmetic unit which executes the programs and coordinates the data transfer within the information system;
- *central memory*, indicated as Random Access Memory (RAM) used for quick storing of data and programs, useful to the processor functioning; its capacity is limited and volatile, i.e. it is lost when the computer is switched off and can contain a small amount of data and programs;
- *secondary memory*, or *mass memory*, used for the durable storing of huge amounts of data and programs; the access to this memory is slower than the central memory;
- the *system bus*, connecting all the functional elements above, enabling the exchange of data;
- *peripheral units*, used to let the computer to communicate with the outside; they include keyboard, cursor and screen terminals, printers, scanning systems and other devices for information transfer from and to the outside, Input/Output (I/O). Each peripheral is controlled by an interface, which lets the processor and the peripheral communicate, lighting the central unit tasks.

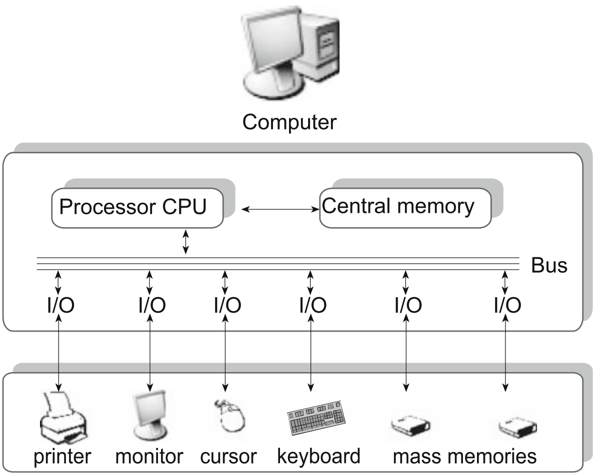


Fig. 5.1 Representation of the physical components of a computer: processor, central memory, mass memory, bus and peripheral devices to communicate to the outside world, for information input and output (I/O)

These components are schematically reported in Fig. 5.1.

Going into more detail, Von Neumann architecture components are constituted by the elements described below.

The processing unit or CPU is composed of

- *Arithmetic–Logic Unit* (ALU), which performs the elementary arithmetical and logic operations.
- *Control Unit*, which coordinates the temporal execution of the functions that are carried out internally to the processing unit or in the other functional elements. The units operate in a coordinated way synchronizing the events through a clock system which scans the operation activation; the clock scanning frequency, or number of activities per unit of time, is measured in Mega Hertz (MHz) and indicates the CPU processing speed. Transfers from one functional element to another are realized by the bus system which performs the logic link.

The other fundamental component of a computer, according to the scheme described by Von Neumann, is the memory. Memories are realized by semiconductor devices and are idealized as big tables whose columns and rows are the memory cells; each cell contains a word. The columns, as many as the word length, correspond to one-memory bit. The information is stored in memory as high- or low-tension state in the memory positions localized by the intersection of rows and columns (Fig. 5.2).

Four main kinds of memory can be distinguished, besides the cache memory (Fig. 5.3):

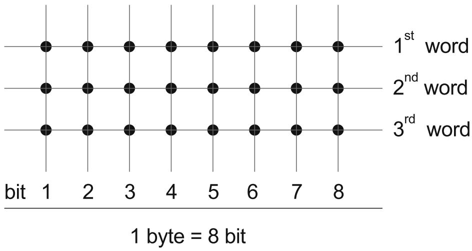


Fig. 5.2 Structure of the memory: the central memory is a sequence of memory cells, each comprehending a word built by a sequence of bit. Eight bit form one byte. Memory is divided up into cells, each cell having a unique address so that the data can be retrieved

- RAM (*Random Access Memory*) or central memory: represents the container of the information that has to be sent to the CPU to be executed. The RAM memory is small and volatile (i.e. the information is lost as the computer is switched off). Both reading and writing operations can be carried out on RAM. The central memory is a sequence of identical memory cells, each containing a word as a bit sequence. Different processors can have different length words. The most used word lengths include the first multiples of the byte (8 bit) and hence there are 8-, 16-, 32- and 64-bit computers. Each cell has its own address, indicating its relative position in order to be found by an address directory in the processing unit.
- *cache memory*: extremely fast memory, built into a computer's CPU or located next to it on a separate chip. Cache memory is RAM that a computer microprocessor can access more quickly than it can access the regular RAM. Although the RAM is fast, it is not enough for modern computers. The solution is to insert, between processor and RAM, a very quick memory to store the most frequently used data (*cache memory*). When the processor loads data from the central mem-

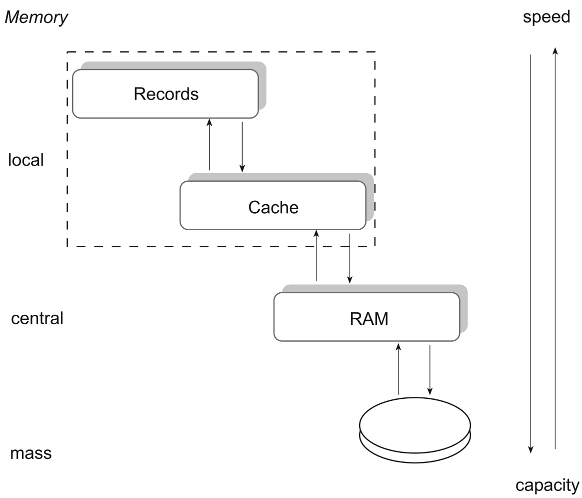


Fig. 5.3 Relationship among the local, central and mass memories in the Von Neumann architecture. The term 'Von Neumann architecture' refers to a computer design model that uses a single storage structure to hold both instructions and data. The term 'Von Neumann machine' can be used to describe such a computer, but it also has other meanings. The separation of storage from the processing unit is implicit in the Von Neumann architecture

ory, data are also loaded to the cache; subsequently the data can be quickly read from the cache instead of from the slower central memory.

- ROM (*Read Only Memory*): not volatile and not writable by the user; it contains data and programs to initialize the system. When the computer is switched on, the RAM is empty. Nevertheless the CPU needs to find from somewhere a sequence of instructions (program) to execute. In each computer is a ROM containing the BIOS (*Basic Input/Output System*). The BIOS is constituted by programs useful to access data that allow the computer to have some basic instruments available when it is switched on. The BIOS contains a small program transferring the operating system (OS) and other essential data from the mass memory to the central one (loading phase or boot).
- *mass memory* (hard disk, floppy disk, CD and DVD-ROM, CD and DVD RW-ReWritable, USB: *Universal Serial Bus*), considered as extent peripherals containing permanent data. The memories can be represented as huge tables where each cell contains information (bit sequences) that can represent data or instructions. The information contained in the mass memory is organized in files, each one with its own name and containing data that can be texts, programs, images.

The relation between the RAM and the mass memory is the same as between the processor cache memory and the RAM. The most used data are first loaded by the RAM and then transferred to the processor.

These quick memories are not used as mass memories because they entail high cost.

The Bus System links the CPU, the memories and the interfaces to peripheral (I/O), mass memory, etc. The Bus is a sub-system that transfers data between computer components inside a computer or between computers. A bus can logically connect several peripherals over the same set of wires. Thanks to this connection system any peripheral inserted in a computer can communicate with the other components, memory and processor included. Modern computer buses can use both parallel and bit-serial connections, and it can be wired in either a multidrop (electrical parallel) or daisy chain topology, or it can be connected by switched hubs, as in the case of USB. The data transfer is under the control of the CPU (Fig. 5.4).

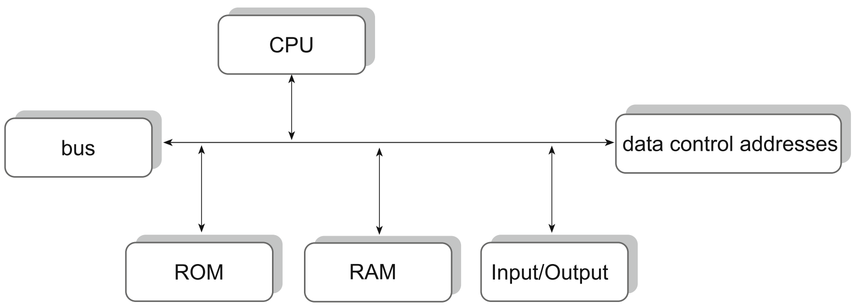


Fig. 5.4 Components of a computer and link networks

When the external environment is not represented by a human user, like in the case of sensors set on aircraft or satellite, the communication is performed by:

- *sensors*, which perceive external phenomena, such as luminous radiation, temperature, eco-radar, measuring physical values or detecting images;
- *carriers*: which translate appropriate commands into actions that can influence the environment all around, such as movements and regulations.

5.2.2.3 Computer Systems

The main characteristic of a general processor is the computation flexibility, as it is based on an architecture not specialized for only one use, but suitable to perform different tasks. Hence a processor is a polyfunctional device, suitable for different applications.

Based on Von Neumann architecture, processors can be distinguished according to their performances and technical characteristics. The main segments into which the computing system is divided are:

- *Personal Digital Assistant (PDA)* or *personal organizer*;
- *personal computers and notebooks*;
- *workstations*;
- *mainframe*: electronic or basic processors.

Chronologically the first to appear were the mainframe, followed by supercomputers in terms of processing capacity, particularly speed of calculation, and by minicomputers. Afterwards workstations were introduced and at last, in the late 1970s, personal computers (PC) entered the market. PCs are nowadays the most used support tool for processing.

Supercomputers could process (October 2004) instructions up to 65 Teraflop, 65 thousand billion operations per second, provided with high speed of data transfer between memory and CPU, working better than scalar processors, which connect in parallel many computers. In a vector processor, by contrast, a single instruction operates simultaneously on multiple data items. In 2008 technology has broken Petaflop speeds, performing 1.105 quadrillion floating-point calculations per second. A scalar processor processes one data item at a time. Breaking the Petaflop barrier, computers enable new types of science unpredictable before. This new generation of petascale machines can move scientific simulation beyond just supporting the two main branches of science, theory and experimentation, and into the foreground.

Minicomputers incorporate most of the functions of a CPU on a single miniaturized electronic circuit.

PDAs are miniaturized multimedia systems having functionalities typical of a personal computer and of a cell phone, with capacity to acquire digital images, to visualize geographic maps, interactive navigation and access to any kind of

information. PDAs directly open opportunities available to a large number of users. They have an important role in the development of the *Universal Mobile Telecommunication System* (UMTS) technology.

The *personal computer*, typically dedicated to only one user, is constituted by a body (containing a central unit, a central memory and a mass memory) connected to a keyboard, cursor, screen and peripherals. The mass memory comprises one or more non-removable disks and removable memories, external to the PC body (optical disks based on laser technology in reading and writing, small-size USB memories simple and quick to use for saving and transferring data between different computers, flash memories, etc.).

Workstations offer higher performance than PCs, especially with respect to graphics and CPU power, memory capacity and multitasking ability.

Big systems, or mainframes, can manage hundreds of users with thousands of terminals, with many computers and huge mass memories.

The capacity of a computer system can enlarge, increasing the memory, adding computers, terminals and peripherals. In particular many computers can be connected to each other forming *computer networks*. Local networks and geographic networks will be described later in this chapter.

5.2.3 Computer Software

The limit between hardware and software is gradual with a neutral part occupied by a micro-programs layer, or firmware, which allows the constructors to store permanent memories on the processor. These programs operate above the hardware layer and can carry out complex functions.

The *machine language* is a programming language operating on the computer's physical components; thus it is complex even for an expert user. In order to overcome these difficulties, higher level programming languages have been created in which each instruction is a letter-and-number string, more easily interpretable by the user, that operates on symbolic components of the machine (relocation). In fact, each electronic processor is equipped with a set of programs that simplify the user–processor interaction level, making the interaction language close to the natural language (Fig. 5.5).

Computer system software can be divided and classified as follows:

- *basic software*, dedicated to processing management, operating above the hardware and firmware; the main program of the basic software is the operating system (OS), managing the processor hardware resources;
- *applicative software*, dedicated to the realization of specific applicative needs, operating above the basic software;
- *personal productivity software*, not requiring programming skills and easy to use for the common user; it includes video-typing systems, electronic notebooks, hypertexts, electronic mail.

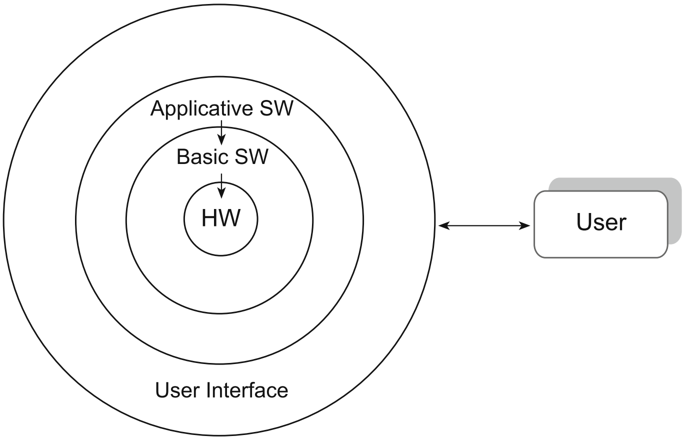


Fig. 5.5 Scheme of the hardware (HW) and software (SW) relationship in a computer and user interface; the SW is an entity contained in the operating system that is contained in the HW. The high-level user interface hides SW and HW to the unfamiliar user

Computer basic software is the operating system, representing the set of programs making the processor operative and usable. It operates over the hardware; the user sees the processor just through an extended or virtual operating system easier to use generated by the OS.

The operating system is responsible for the management and coordination of activities and the sharing of the resources of the computer. The OS acts as a host for application programs that are run on the machine. As a host, one of its purposes is to handle the details of the operation of the hardware.

The OS has two main functions:

- to interpret and execute elementary commands;
- to organize the mass memory structure.

It is responsible for the communication among computers, essential functions in modern network architectures.

In a multi-user system, the OS has to perform the functions of manager of the available resources:

- processing unit, launching the execution of the required programs;
- central memory, ensuring the required programs and data are loaded into memory for a dedicated service to each user without interfering with the others connected.

The common functions and the services available to the OS are used by the set of programs called *applicative software* as they are dedicated not to managing the processor but to solving the problems due to a defined application. Among the most diffused, which represent only a small part, are

- DataBase Management System (DBMS);
- spreadsheets or automatic computation sheets;
- word processor for text processing;
- graphics programs which can be finalized to photographic image processing, planning (*Computer-Aided Design*, CAD), illustrating, presentation, etc.;
- programs for vector data visualization and processing, and others.

Applicative programs are above the operating systems and use high-level programming languages. These programs are not influenced by the characteristics of the system architecture and their transfer to another computer system can be easily realized.

DBMSs provide a software layer which facilitates access and data sharing by different functionalities typical of the applicative programs and thus can also be classified among the basic software systems.

The software is also responsible for data transmission management and for the computer networks. The transmission channels carry electrical signals, while communication protocols, or communication software, assure the correct transmission of data from the terminals to the processor or vice versa, or between two processors connected in the network.

Due to the importance of the choice of the OS, in relation to stability, dependence on the producers and cost, and due to the relevance of the DBMSs within the domain of geographical information, these two aspects are analysed in more detail.

5.2.3.1 Operating Systems

An operating system is the software which, once loaded into memory and executed as the computer is switched on, allows the applicative programs to run and to use the processor hardware resources. The OS dictates to the user the way to operate the processor, establishing what the user can and cannot do and the order in which some operations must be executed (Fig. 5.6).

DOS

The *Disk Operation System* (DOS) is the operating system upon which the history of the computer was built.

In 1979 the American company *Seattle Computer Products* realized a card with the 8086 processor for the S-100 bus. One of the first of these cards was available to Microsoft, a small company producing languages started by Bill Gates and Paul Allen, known after developing the programming software *BASIC for CP/M*. Two weeks later, at the computers national conference in New York, they showed a version of BASIC for 8086 processors: the DOS.

DOS was planned to run on a personal computer, a processor used by single users. DOS is also conceived to execute one program at a time: while the editor is used, it is not possible to use a communication program or execute a utility program.

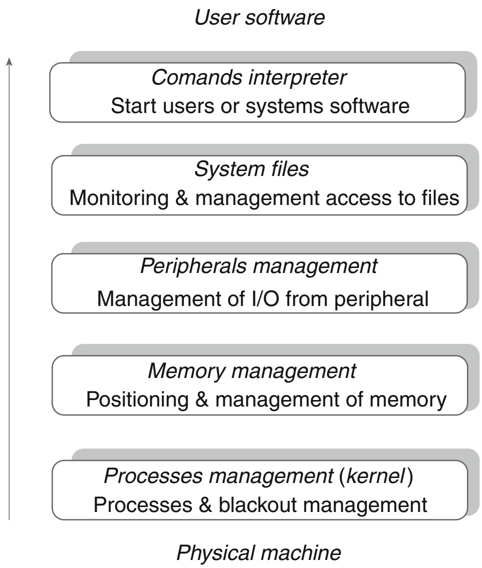


Fig. 5.6 Operating system's stratified architecture, bottom-up; each layer represents a virtual machine

This limit in practice determined the end of DOS. Other important limits regard the use of memory and of graphics.

The following is an example of a DOS instruction:

C:\DOS\>DIR *.txt >lista.txt

MacOS

On January 24, 1984 a new conception processor was introduced, the *Macintosh* by Apple, an American company. It was an original processor at that time, both for its hardware structure and the operating system, the MacOS (*Macintosh Operating System*). The purpose was to produce a processor that not only would be used by experts but would provide instruments for a quick and widespread use. The main characteristics were in accordance with the following basic rules:

- compatibility with other software and peripherals;
- easy installation of programs;
- reduced necessity of training courses as all the programs work with the same logic, the same commands, have the same aspect and use the same system resources. Moreover the representation on the screen is identical to the result obtained by printing;
- intuitive operating system, visual and with commands not requiring any syntax.

Until that time computers used to work in text format, by means of graphics only for particular purposes. With MacOS the text format was eliminated and visualization with a graphical method was introduced.

Windows

Windows owned by *Microsoft* manages the hardware and software resources of a processing system, allowing the user to refer to these resources by symbolic commands or graphic representations, offering a graphic interface that can be used by everybody and much easier to use than the textual interface typical of previous programs, like DOS.

Windows modules include the main functions of an operating system; Windows also includes modules to manage local networks, the Internet, multimedia tools.

The graphic user interface facilitates both the management of the hardware components of the system (peripheral driver) and the management of the software and of the archives of data stored on different supports (magnetic and optical disks, USB, etc.) and makes them autonomous.

Windows is completely integrated with the access to the Internet network, providing a unique visualization for the local and provided by the network resources.

The Windows system concretizes the metaphor of the desktop by the graphic interface:

- the screen represents the working plane;
- the icons, or images, represent the objects used in the work;
- the windows represent the paper sheets that can be added or removed from the desktop. A window is a mobile object on the screen and follows, in the position and size, the operator commands.

The icons represent the tools that can be launched by the user; the windows are the active programs; the tool bar indicates the open windows on the desktop and can be personalized by the user. The OS used by the graphic interface is based on the cursor positioning device (mouse), simple to control by pressure on the button images and documents on the screen. The inclusion of a tool bar enables the user to remember previous actions by a simple pointing on the application bar. This tool introduces the possibility to activate multitasking processes, an important characteristic of the modern operating systems.

UNIX

A processor operated by UNIX is conceived in a completely different way than a personal computer (PC). There are also several versions of the UNIX operating system for PCs, but they are little used.

UNIX is much more than a powerful DOS, more flexible, but also more difficult to learn; nevertheless working with this powerful instrument opens wide possibilities in comparison with DOS's narrow environment.

Unix OS was created around the early 1970s at the same time as the first minicomputers (new processors more compact and quicker than basic big electronic processors (mainframes) produced until that time) were introduced. UNIX is a compact and modular operating system, adaptable to the hardware resources and to the operative needs of an operative environment, characteristics which quickly determined its huge success. As UNIX and the minicomputers were introduced, the use of alphanumerical terminals multiplied, which allowed easier and more efficient communication between the user and the system (until that time perforated cards had been used). Also, the concept of an open system and of distributed processing started to be used. UNIX communicates with the other systems and to be interconnected with other processors in order to operate on resources not centralized in one large system, but distributed in several medium power processors.

UNIX is based on a central nucleus (kernel) managing low-level communication with the processor, program execution and sharing or protection of system resources.

The reference standard is the Berkeley UNIX or *Berkeley Software Distribution* (BSD). It is the operating system developed and distributed by the *Computer Systems Research Group* (CSRG) of the University of California, Berkeley, from 1977 to 1995. Today, the term BSD is non-specifically used to refer to any of BSD's descendants, e.g. FreeBSD, NetBSD, OpenBSD, which together form a branch of the family of UNIX-like operating systems. The UNIX System V, with its version SVR4, is probably more common than BSD.

UNIX is an operative synthetic and essential system. It is versatile and ductile, never redundant.

The fundamental characteristic of UNIX is to be a multi-user operating system: many people can use the system at the same time or in different moments when the resources are shared by many people. Any operation that might overload the system or slow down another user is carefully avoided, in order to leave as many resources as possible available to the real program processing.

UNIX executes many programs at the same time; for example if the system is asked to execute two processes, A and B, the *Control Processing Unit* (CPU) executes alternately process A and process B. The OS controls that the CPU equally carries on all the active processes; when an error occurs, the other processes and the system do not suffer any damage and can continue working without any negative effect on their tasks.

Example of UNIX instruction:

```
drwxr-xr-x  2   dosuser dos   256   Apr  28  15:40 ./

drwxr-xr-x  15  dosuser dos   928   Apr  28  14:50 ../

-rw-r–r–    1   dosuser dos   20    Apr  28  16:05 .bash_history
```

Linux

Linux is an Open Source operating system available for several hardware platforms. It belongs to the family of UNIX systems (Solaris, AIX, HPUX, SCO, etc.) but it has been conceived to be compatible with the characteristics of other operating systems.

Linux itself is constituted by one OS kernel which controls the functioning of the processor. This nucleus was ideated by Linus Torvalds, a Finnish student at the University of Helsinki, in 1991. Linux's kernel and its software are distributed under the *General Public Licence* (GPL) or the *GPL library* (LGPL). These licences ensure both the copyright by the author and the possibility of modifying the source code by the software user.

The GPL licence was the starting point of the movement for the Free Software used by thousands of programmers spread all over the world who contribute to Linux's development. The work is coordinated by the use of the Internet. The kernel sources are available on the network both in the stable version, which ends with an even number, and in the development version, ending with an odd number.

As regards Linux, there is a very wide availability of applicative software tools available in many archives on the Internet.

Catalogues of the software are available both in console and graphic mode, from which it can be downloaded as a source form to be compiled, or ready to be used.

5.2.3.2 Databases

At the end of the 1960s, many researchers started to face the separation between the logical and the physical model and to consider the data issue in an integrated way. A new concept, the database, emerged from these studies.

A database is a structured collection of organized data. A computer database allows inserting, modifying, cancelling, organizing and restoring data.

The software models the database structure in database models. Geographical and alpha-numerical data conceptual organization in databases is characterized by the structure of the management system called *DataBase Management System*, DBMS (Fig. 5.7).

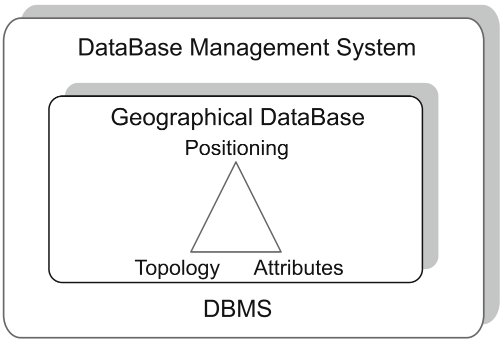


Fig. 5.7 Geographic database and database management system (DBMS) relationship

DBMSs are complex sets of software programs used to organize, store, manage, retrieve and maintain data in databases. These are categorized according to the database model that they support. The model tends to determine the query languages that are available to access the database. An important aspect of the internal engineering of a DBMS, independent of the data model, is concerned with managing factors such as performance, concurrency, integrity and recovery from hardware failures.

A DBMS includes

- modelling language to define the organization of each database hosted in the DBMS, according to the data model;
- data structure, as fields, records, files and objects, optimized to deal with very large amounts of data stored on a permanent data storage device;
- database query language and report writer to allow users to interactively interrogate the database, analyse its data and update it according to the users' access to data. It also controls the security of the database. Data security prevents unauthorized users from viewing or updating the database. Query languages are computer languages used to make queries into databases and GIS. They can be to a large extent classified according to whether they are database query language or information retrieval query language;
- transaction mechanism, that ideally would guarantee *Atomicity, Consistency, Isolation, Durability* (ACID), a set of properties that guarantee that database transactions are processed reliably. A transaction, in the context of databases, is a single logical operation on the data. A transaction mechanism ensures and maintains data integrity, despite concurrent user access and faults.

The DBMS accepts requests for data from the application program and instructs the operating system to transfer the appropriate data.

The most used DBMS models, differentiated according to the data archiving systems that they adopt, are the *hierarchical,* network, *relational* and *object-oriented* models.

- *hierarchical model*: with data organized in a tree structure; the top of the hierarchy is represented by the root which has one or lower elements related to it (Fig. 5.8). This structure arranges the data elements in a hierarchy and helps to establish logical relationships among data elements of multiple files. Each unit in the model is a record which is also known as a node. In this model, each record on one level can be related to multiple records on the next lower level excluding connection among the elements at the same level. The rigid structure is disadvantageous in accommodating the dynamic needs of an organization of data not strictly hierarchical.

- *network model*: similar to the hierarchical ones although more flexible as connections among elements belonging to the same level are possible (Fig. 5.9a); The network model tends to store records with links to other records. In network DBMS, each element or group of similar elements is connected to different ele-

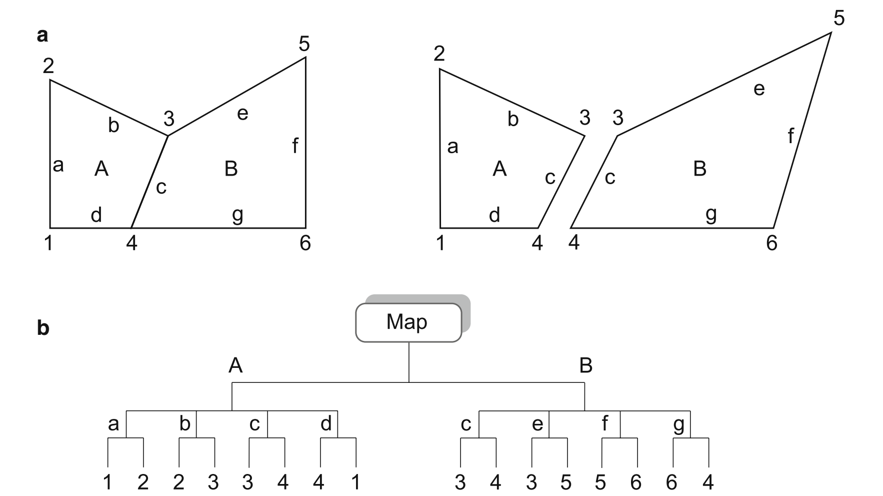


Fig. 5.8 (**a**) Map attributes for two polygons; (**b**) DBMS with hierarchical data structure

ments at different levels. The structure elements can be connected by relation: one to many, many to one and many to many;

- *relational model*: data are organized in tables where the rows correspond to the *records* and the columns to the *fields* (Fig. 5.9b); in this structure, there is not a hierarchy of fields within the records: each field can be used as a research key through the queries. Data are stored as sets of values in the records and grouped in bi-dimensional tables, each one saved as a single file. The tables in a relational database respect three basic rules.

  - the ordering of columns is immaterial;
  - identical rows are not allowed in a table;
  - each row has a single value for each of its columns.

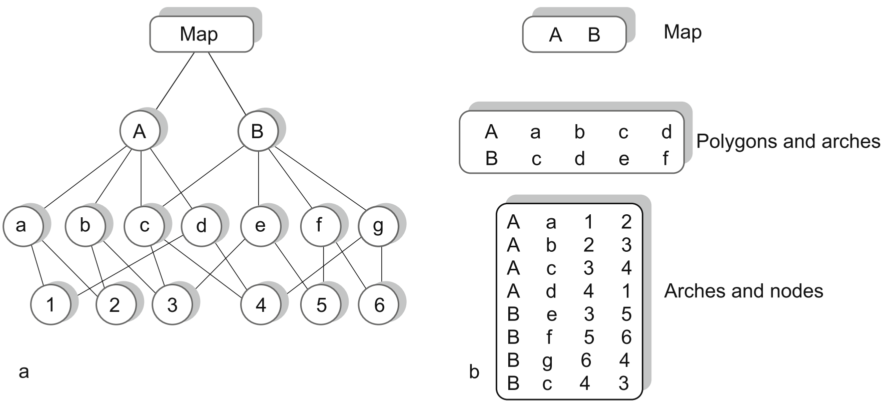


Fig. 5.9 (**a**) DBMS with network linkage data structure; (**b**) relational data structure with reference to the polygons in Fig. 5.8a

If the same value occurs in two different records, it can imply a relationship between those records.

The most used query language in the relational model is the *Structured Query Language* (SQL), standard for industrial software. It is composed of

- DDL (*Data Definition Language*) for the database definition;
- DML (*Data Manipulation Language*) for the database manipulation and query;
- DCL (*Data Control Language*) used to control access to data in a database.

The relational model is the most commonly used in GIS (Fig. 5.10); it gives space to any possible connection, while in hierarchical and network models the connections have to be known a priori;

- *object-oriented databases*: the most recently developed and experimented. In this model, any complexity and structure model can be precisely represented as a homogeneous independent entity. The basic elements in modelling are objects and classes.

The planning of a database is a long but necessary process; although it represents only an aspect of the more general problem concerning the software planning, data organization is fundamental for spatial information management (Fig. 5.11).

A database realization requires a characteristics acquisition phase. Data, limits and operations that constitute the database have to be specified and described.

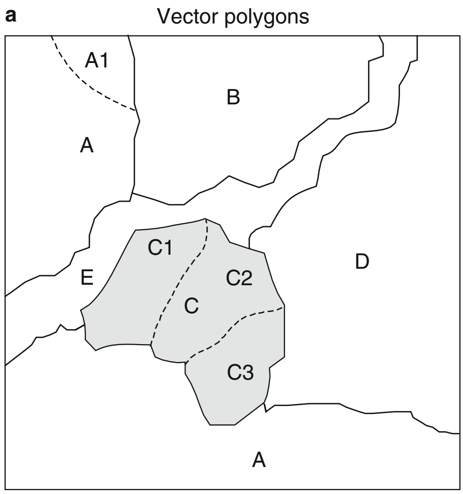
The characteristics are provided by the available documents, by data management of previous systems and by the users. The result is a natural language not ambiguous text that describes the problem and assesses the solution's complexity and cost.

Database planning requires two main phases (Fig. 5.12):

- *conceptual planning* with the definition of a data model in a textual or graphic formal language. This phase consists in transforming informal details of the reality to be represented into formal descriptions;
- *technological planning* and relative implementation by the use of a DBMS. This phase requires a further representation of the information through a new model (for example the relational model based on data organization through tables, one connected to the other). The DBMS enables the user to operate by access procedure on the data at logic level (data structure), independently of their physical representation (data).

A data model is defined by the abstraction of the reality into concepts representing the reality itself. A conceptual data model is the whole of symbolic structures representing the abstract concepts independently of any processor. A logic data model translates the conceptual structures into logic structures that can be processed by a DBMS.

A database transformation from conceptual model into logic model passes through a DBMS.



b

Landscape unit	Geomorphology	Land use	Land unit
A	plateau	Bare soil	A1
B	caldera	cultivated	B5
C	glacis	cultivated	C3
D	caldera	cultivated	D2
E	escarpment	pasture	E4

Fig. 5.10 Networking in a relational database between (**a**) geographical data as vector polygons, defined by cap letters and numbers and (**b**) alpha-numerical data. The example refers to a soil erosion scenario

A DBMS enables the user to operate on the data themselves at logic level, independently of their physical representation. The DBMS can

- integrate data, avoiding inconsistencies;
- regulate the access to the data, assuring confidentiality and authenticity;
- manage the competition in the case of distributed databases;
- restore the data in case of error.

There are two main categories of database:

- *personal*;
- *client/server*.

In case of personal database applications, DBMS are conceived for individual users. The personal programs and databases are stored in the same processor. When

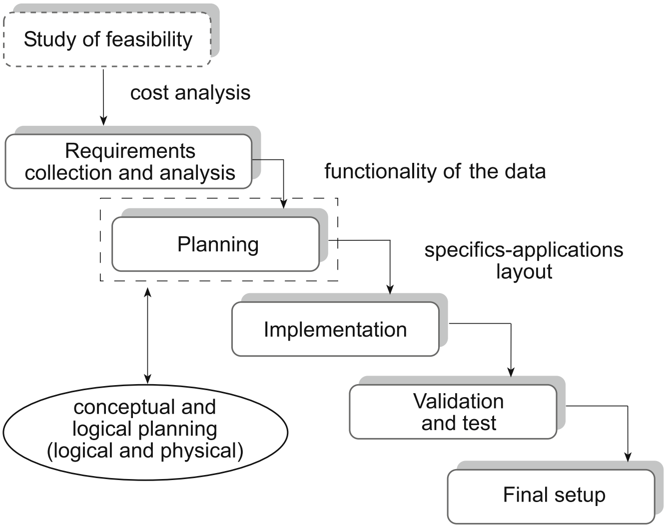


Fig. 5.11 Steps in a database redaction

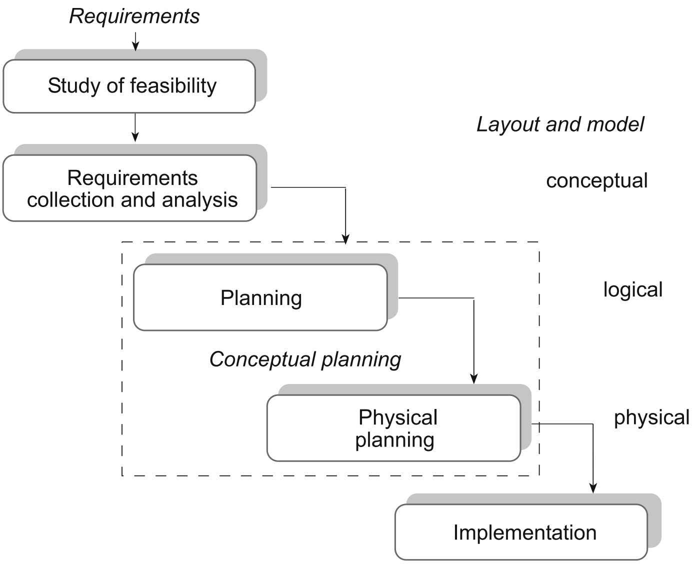


Fig. 5.12 Scheme of the design of a database in the conceptual and technological phases

a user launches the query to examine, insert, update or eliminate data, the query requires finding all the data in the databases, in the DBMS software and in the programs to process the data.

Client/server databases, instead, separate the DBMS from the programs that operate on the database. The DBMS is installed and executed on a *server*, while the

programs using the database are installed and executed on a *client*. The programs executed in the client send queries to the DBMS through the network. When the client/server DBMS receives a query, it executes the required functions about the data (finding, ordering, filtering, etc.) and gives back to the client, through the network, only the final result of the request (Fig. 5.16).

In a personal database system, when a working station interrupts the procedures due to a software error or to an electric power interruption, damage can occur in the database due to interrupted command queries. A partially written record, in a database file, can sometimes make the whole file unreadable. The command queries in execution during the fault generally cannot be reconstructed. If it is not possible to repair them, the database administrator can restore the last saved version, but insertion, modification and cancelled operations which occurred after the last saving are lost.

If the queries of a client station fault do not affect the other users. The larger part of the client/server DBMS server offers many additional characteristics to minimize the possibility of faults and, when they occur, these servers have quick and efficient restoration mechanisms.

5.2.3.3 Database Management Software

There are several types of database management software available, conceived for the hierarchical, network, relational and object-oriented models. Two of the most common are

- *Relational Database Management System* (RDBMS) is a database management (DBMS) based on the relational model. The system, defined in 12 fundamental rules, has been introduced by Codd in 1970. The term now describes a broader class of database systems. At the minimum, these systems present the data to the user as relations in tabular form and provide relational operators to manipulate the data in the tables.
- *Object-Relational DataBase* (ORD) or *Object-Relational DataBase Management System* (ORDBMS): database management system similar to a relational database, but with an object-oriented database model where objects, classes and properties are directly supported in database configuration and in the query language.

The aspects concerning data spatial analysis in GIS are reported in Chapter 9.

5.3 Network Architecture

The necessity to transfer data and information at a distance, diffused and accessible to everybody, stimulated the conditions to develop computer networks. The most known form is the Internet, the computer network allowing everybody to get connected with everyone else from any part of the world.

The distributed systems, or computer networks, are constituted by many processors able to interact and communicate among them by appropriate peripherals, and which share resources like data and hardware. Computer networks link nodes and provide several network services to the connected computers. These are the product of the cooperation between computer science and telecommunication.

The advantages of participating to a computer network are

- exchange of data, information and programs between processors and terminals located in different places;
- possibility of sharing resources installed on a processor with other computers;
- resolution of anomalies or damage if an OS is out of order, using another system of the network to replace it (back up functions) without performance decrease;
- mobile connection to the network by portable terminal devices.

5.3.1 Transmission Mode

Data transmission allows two computers to be connected or a computer to connect to many remote terminals.

Data transfer is characterized by the *speed of transmission*, measured in bit per second (bps).

The available means mostly used for transmission are

- *Unshielded Twisted-Pair* (UTP) copper wiring commonly used in telephone cables for communications, also allowing data transmission across long distances, with limited speeds, 9600 bps max;
- *co-axial cable*: ideated to converge the television signal; it is an electrical cable consisting of an inner conductor surrounded by an insulating spacer, surrounded by an outer cylindrical conductor. It is appropriate for local connection quick data transfer (10^7 bps);
- *optical fibre*: for systems distributed across long distances ideated for quick data transmission (10^9 bps). It is a glass fibre that carries light along its length. The signal is transmitted by photoelectrical diodes that emit light, codifying the binary numbers as the presence or the absence of luminous signal;
- *EM waves*, or wireless networks: for short-distance microwaves or infrared transmission, or via satellite long distance radio waves in specific transmission frequency, as the signal moves at light speed.

Telephonic lines were born to transport an analogical continuous signal, i.e. the voice. The data produced and distributed by a computer in binary form are transmitted as digital signals characterized by a squared shape wave. Devices that can transform analogical signals into digital ones and vice versa have been developed. A digital signal has to be converted into an appropriate analogical signal before

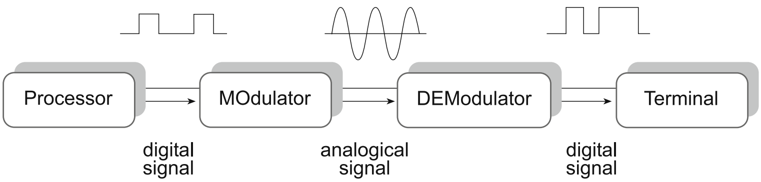


Fig. 5.13 Analogical–digital modulation and demodulation process (MODEM) of the signal for data transmission across the telephonic network

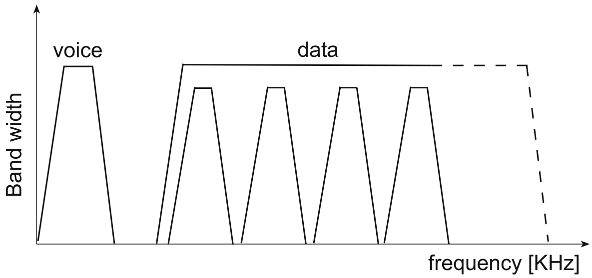


Fig. 5.14 Scheme of a larger band use with respect to the traditional transmission

being transmitted along the telephone line and has then to be converted back into a digital signal at the end of the transmission.

Devices for the digital transmission of data can be summarized as follows:

- *modem*, acronym for *MOdulation and DEModulation*, able to convert the digital signal into analogical signal for data transmission along copper telephonic lines, which allow the passage only of analogical continuous electric signals (Fig. 5.13);
- ADSL (*Asymmetric Digital Subscriber Line*), data communications technology that enables fast data transmission over telephone lines. A splitter allows a single telephone connection to be used for both ADSL service and voice calls at the same time. ADSL uses a high-frequency band to carry data, images and audio in numerical format (Fig. 5.14). ADSL connection is permanent; it is not necessary to get connected to the Internet through an *Internet Protocol* (IP).

5.3.2 Network Logical Scheme

The minimum network that can be realized is the connection between two processors connected by a transmission line (Fig. 5.15); each processing unit is provided with:

- an operating system;
- applicative software (AP) distributing services to local users;

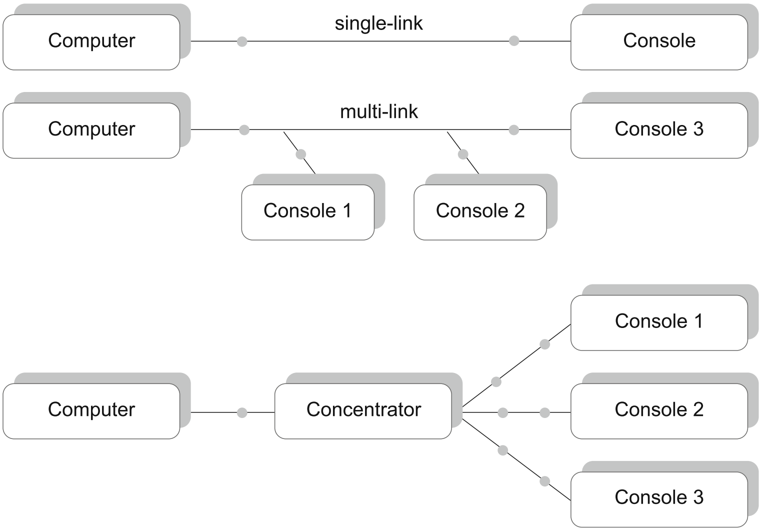


Fig. 5.15 Three possible terminal connections

- data acquisition/distribution devices (input/output, I/O);
- a device for data transmission, called *Data Terminal Equipment* (DTE).

The *client station* launches a specific request (A); the *server* supplies (B), making its data and programs available (Fig. 5.16).

The same physical connection can be used to support two different logic interactions at the same time:

- the *logic connection*, i.e. the formulation of a request by one of the two units and its correct interpretation by the other for the elaboration of the reply;
- the *physical connection*, constituted by the connection line through which data and information move.

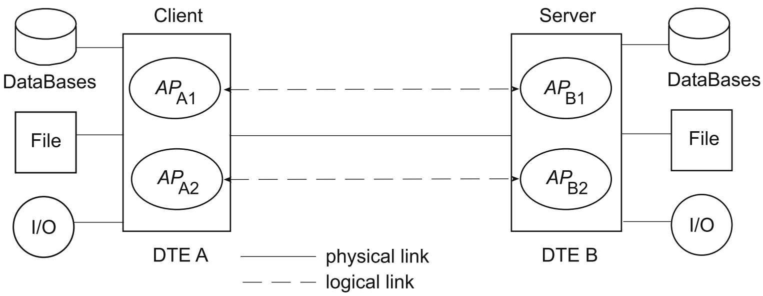


Fig. 5.16 Client–server connection: AP: applicative software; DTE: data terminal equipment

Two or more DTEs' communicative interfaces dialogue by using protocols.

There are standard limits to the highest number of computers that can be connected to a multipoint network configuration in order to avoid a degradation of the transmitted signal.

5.3.3 Network Typology and Digital Transmission (Classification WAN/LAN)

Computer networks are divided into two broad categories:

- *Wide Area Network* (WAN), or geographic networks, worldwide extended;
- *Local Area Network* (LAN), or local networks.

WAN's main characteristics are

- very high number of computers connected (of the order of 10^6);
- public management entities which rent the use of communication lines (telephone or optical lines) to the users;
- topological mesh structure to distribute the traffic over many paths;
- transmission through intermediate nodes;
- control centres displacement, in order to verify the accuracy of the transactions, and commutation nodes (*Data Switching Equipment*, DSE), to canalize the traffic.

The typical characteristics of a LAN, differing from a WAN, are

- LAN extension is geographically limited;
- the channels used for the transmission are private;
- data transmission is quicker than in a WAN;
- they are reliable and with limited errors in the transmissions;
- topological configurations adopted for LANs are ring, bus and star (Fig. 5.17).

5.3.3.1 Local Network Configuration (LAN)

The management of business networks with security and privacy needs is a problem that requires careful study and planning by specialists. If the purpose is just to connect a small number (2, 3, ...) of computers, with limited needs as sharing files, printers, scanning systems, and other common peripherals, the procedure is easier. In a typical LAN configuration, one computer is designated as the file server. It stores all of the software that controls the network, as well as the software that can be shared by the computers connected to the network. Computers connected to the file server are called *client*. The clients can be less powerful than the file server, and they may have additional software on their hard drives. On most LANs, cables are used to connect the network interface cards in each computer.

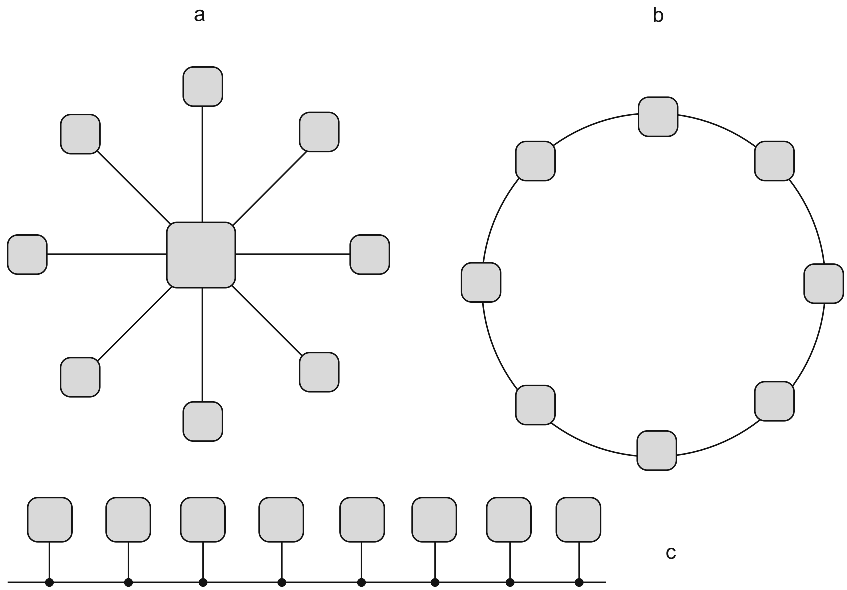


Fig. 5.17 Topological categories: (**a**) star; (**b**) ring; (**c**) bus

The terms to know for a LAN configuration are

- *server*: the most important computer in the network, allows the processors to use the shared resources and has administration functions; for this relevance, the processor chosen for this function is the most powerful processor among those available;
- *client*: any PC connected to the network not as a server;
- *network board*: it makes the PC to communicate with the other components of the network;
- *HUB*: device that collects all the PCs connection cables, to which all the computers of the LAN are connected;
- *cables*: they physically connect the processors of the network;
- *communication protocol*: the common language used by the computers to share data.

To build a LAN, a specific hardware is required:

- a network board for each PC that is part of the LAN;
- a HUB connecting the PCs.

The HUB speed has to suit the network board's speed and needs a sufficient number of ports to connect the processors that constitute the network.

5.3.4 Network Topological Relationships

Networks can be divided into the following main topological categories based on the different functional links among their components (Fig. 5.17):

- *tree network* with many hierarchical levels, the highest of which are occupied by the most powerful system of the structure. The management is simple, but there are some inconveniences in relation to the absolute dependence of the network on the top level;
- *star network*, where the peripheral nodes are connected to one main node independently from the others. Thus all the communications pass through the central node;
- *ring network*, when all the nodes are connected among them in sequence, in order to form an ideal ring, where each node has a direct contact only with the previous and the following;
- *bus network*, where the node connection is shared by all the nodes, so the data transmitted by one node can be intercepted by all the others. The simultaneous transmission by two nodes generates a collision and the loss of the transmitted message;
- *mesh network*, where the nodes are connected by several circuits. This network configuration is reliable for the existence of multiple paths between two stations offering good performance in terms of transition speed.

5.3.5 Communication Protocols

A communication protocol is the set of conventional rules for data representation, signalling, authentication and error detection required to send information over a communication channel.

The transmission across a line connecting two stations needs a line protocol which regulates the flux. There are different types of line protocols each of which, independently of its specific characteristics, accomplishes the following tasks:

- synchronization of the transmitting and receiving station, before and during the transmission;
- exchange of data with a formal structure in order to differentiate the control information that goes with them;
- management of the flux and the exchanges;
- control of the transmission errors and, if possible, their correction.

Line protocols can be distinguished as follows:

- *primary/secondary* relation (master/slave) typical of the hierarchical structure multipoint lines, in which a processor plays the role of primary station, while the lower level processors of the same line are secondary. The primary station controls and manages the others as regards both the functionality and the transmission, defining the priority rules in the use of the line;
- *peer-to-peer* protocols, according to which all the stations connected to the same line have the same rights for transmission, without a primary station role. Peer-to-peer protocols are used mainly in LANs;

- *hybrid line* protocols, which can operate in both modes, i.e. peer-to-peer as well as related primary/secondary.

5.4 Network Infrastructures

5.4.1 Internet

The unifying principle implicitly accepted by every user of the network is the one of reciprocal exchange. By connecting their own local network to the Internet network, all the other users are allowed to access the network public domain services and to go through the network to connect to other local systems.

The Internet is the most important geographic network worldwide and can be defined as the whole of the computers and networks, located all over the world, which are physically connected in different ways. The Internet is the *network of the networks*, namely the result of a very high number of LANs spread across the world and interconnected. It is a universal communication tool, a *computer highway* able to transport any type of information on the condition that it is codified in binary form, in continuous evolution.

The processors in the network are interconnected according to a functional hierarchical model, namely the communication protocols (Internet protocol), of two types:

- *low level*, if they are in charge of binary data transmission along the network, without caring about their real meaning: the *Transmission Control Protocol* (TCP) and *Internet Protocol* (IP), ideated in 1974 and network standard since 1983;
- *high level*, if they carry out more complex operations mainly consisting in the codification of data to be transmitted and in the decodification of the received data, independently of the way in which they are moved through the network. This level includes the protocols *File Transfer Protocol* (FTP), *HyperText Transfer Protocol* (HTTP), *TelNet*, *Simple Mail Transfer Protocol* (SMTP), etc.

A connection to the Internet requires a dedicated service through a provider. An *Internet access provider* (IAP) is a company or business that provides access to the Internet and related services.

5.4.2 World Wide Web

The *World Wide Web,* indicated with the short form WWW or W3, is the widest set of informatics objects available on the Internet.

The main characteristic of Web objects, usually called hyper-objects, is to contain, besides a set of static information, many links or *hyperlinks* to other hyper-

objects, realizing a Web of virtual links. The most diffused hyper-object is the *hypertext*, a text type electronic document which in many points recalls other objects. Current Web technology connects texts, data, images and sounds through hyperlinks and downloading them on the local disk. Born in the CERN (*European Organization for Nuclear Research*) Geneva as a tool to easily exchange complex structure information all over the world, it became the main Internet communication tool thanks to two fundamental properties of the hyper-objects:

- dynamic refreshing of information by the user;
- independent access from the physical location and of the hardware device, virtually eliminating the geographic distances and the differences among the computers connected to the network.

The World Wide Web (WWW) represents the main aspect of Internet, with which it is often confused:

- Internet is a huge computer network that physically exists in the form of cables, electronic devices and hardware;
- the WWW is a set of digital resources interlinked that run over the network.

The virtual links, generically called hypertextual, are supported by high-level communication protocols which allow the transferring of hyper-objects through the network, properly codifying and decoding the programming language with which they are built or described (*HyperText Markup Language, HTML*). The *HTTP* represents the standard method for Web information exchange through the definition of the rules about the dialogue between the client browser and the enquired servers.

HTML is the markup language for Web pages. It provides a means to describe the structure of text-based information in a document and it includes virtual links with any other type of objects.

Hypertexts are *ASCII files* that can be created or modified by any editor. They are different from a normal document by the presence of instructions or textual markers (*tags*) enclosed in angular brackets, defining the text formatting and the insertion of images, sounds, animations or hypertextual links to any other hyper-object. A hypertext is constituted by the following elements:

- contents, data and information that are supposed to be distributed through the network;
- tags or commands, to indicate to the browser how to represent the different contents;
- any comment that, enclosed in particular tags, are the support instrument for updating or modifying the Web page in the HTML codification (e.g. <center> </center>).

The contents have to be enclosed between pairs of similar tags that define the representation typologies and the modes to mark the beginning and the end of each object of the hypertext; the final tag is different only for one more slash (/).

The pair <html>...</html> includes the whole document, while <head>...<head> includes the header, i.e. the information that the browser internally uses to characterize the hypertext (for example, the title) without visualizing them. The tags <body>...</body> instead contain the hypertext body.

An HTML document has the following basic scheme:

```
<html>
<head>............</head>
<body>.............</body>
</html>
```

HTML has complex animation potentialities and a more efficient management of the applications (scripts) that can be launched from a Web page. The scripts can be divided into two categories:

- *client side*, creating interactive pages, as, the users can modify the content of the pages that are addressed from the client to the server;
- *server side*, creating pages to be sent to the client exactly as he requests, also in relation to the information provided by the user himself. Among the languages, there are *Personal Home Page* (PHP), Perl, *servlet Java* and *Java Server Pages* (JSP) and the *Active Server Pages* (ASP) technology.

5.4.3 WWW Navigation

The basic principle of the World Wide Web (WWW) is the free access, share and globalization of the information. Besides an appropriate pseudo-programming language (HTML) and a protocol for their transfer (HTTP), also some programs that permit the visualization their contents are necessary. These applicative software tools are called *browsers* and allow to read, open or download the hyper-objects from the WWW on the local disk, but not to modify them.

The movement through the World Wide Web (also called *netsurfing* or *surfing*), namely the passing from one Web resource to another, follows a standard procedure independent from the used browser:

- the client browser reads the URL (*Uniform Resource Locator*, WWW address) of the object to be accessed and forwards a request to the correspondent Web server using HTTP;
- using the same protocol, the server replies to the request sending to the client the object to which the URL referred;
- the browser reads and interprets the HTML codes that constitute this object, providing to the user in the correct format.

5.4.4 Intranet

Intranet is a local, private network realized by the application of Internet high-level standards and protocols. The users of this network can securely share the network connectivity, the languages, the standards and the programs (like the browsers) created for remote communication.

Intranet access is provided throughout a gateway with a firewall, along with user authentication, encryption of messages, and often makes use of *Virtual Private Networks* (VPNs). Through such devices and systems, off-site authorized users can access computing resources, information and internal communications.

To reduce the network traffic and improve network performance, different strategies in network management can be adopted. Networks can be structured according to different organization models:

- *centralized*, with servers administrated and controlled in a centralized way;
- *decentralized*, the users are scattered in different locations, all feeding information into the content management system;
- *mixed*, when a centralized control is set for just some of the available resources.

Better network performance results if the network traffic is kept as local as possible.

5.4.5 Network Security

The Internet is first of all a network of people who communicate among themselves. The conversation can be public domain, and in this case the network is a means of universal distribution, or can be private conversation.

The input data go from a node of the network through many cables and are treated by a certain number of computers before reaching the destination. If hackers wanted to intercept the communications, they could decode and read the content, knowing that the IP addresses both of the sender and the receiver are indicated in each package. There are programs called sniffers which can intercept the messages going through the local network and select the desired ones based on the sender, the receiver or the content. In order to avoid intrusions before accessing a computer or a specific resource (e.g. a Web site), the execution of a *login* will be required by the insertion of a *username* and a *password*. This system only assures a medium–low level security.

To better protect the output reserved messages, it is possible to adopt cryptography systems which use transformation algorithms (*key*) to convert the message's text (*plaintext*) into *ciphertex*.

5.4.6 Wireless

WiMAX (*Worldwide Interoperability for Microwave Access*) technology allows access to broadband telecommunication network without wires (*Broadband Wireless Access*, BWA).

The acronym has been defined by WiMAX Forum, a consortium, founded in 2001, of 420 Enterprises with the scope of developing, monitoring, promoting and testing the systems' interoperability based on the WirelessMAN (*Wireless Metropolitan Area Network*) standard.

The main characteristics of WiMAX that should contribute to the diffusion of this competitive standard are

- *flexibility:* WiMAX is able to support point to multipoint systems (P–MP) and multipoint to multipoint systems (MP–MP), called also *mesh*;
- *security*: WiMAX implements several techniques of encryption, security and authentication against hackers and intentional intruders.
- *Quality of the Service (QoS)*: WIMAX supports five typologies of QoS:
  - *Unsolicited Grant Service* (UGS): real-time systems with fixed dimension (i.e. VoIP, Voice over Internet Protocol);
  - *real-time Polling Service* (rtPS): real-time systems with variable size data (i.e. video application);
  - *non-real-time Polling Service* (nrtPS): data flux tolerance of delay. The service offers unicast polls on a regular basis, which assures that the flow receives request opportunities even during network congestion. (i.e. FTP, *File Transfer Protocol*, application);
  - *Extended Real-Time Polling Service* (ErtPS) similar to rtPS for real-time flux with fixed dimension (e.g. VoIP with silence suppression);
  - *Best Effort* (BE): data flux where a minimum service level is not requested.
- *throughput*: WiMAX transmits a huge amount of data with high efficiency, more than twice the capability of HIPERLAN (*HIgh PERformance Radio LAN*) a European WLAN standard, able to guarantee a throughput of 24 Mb/s on frequencies of 2.4 gigahertz;
- *installation*: WiMAX requests only a small antenna to be operative;
- *interoperability*: WiMAX is an independent standard free from any constriction deriving from devices and providers;
- *mobility*: WiMAX allows mobile connections up to 160 km/h;
- *costs/cover*: the open standard and the economy for the large-scale diffusion of the system should reduce drastically the costs to the end user, giving the cover between *Base Station* (BS) and *Subscriber Station* (SS) and high-speed data transmission;
- *Not Line Of Sight* (*NLOS*): capacity of transmission throughout impeding objects, as complex orographic landscapes, tree tops or artificial infrastructures, or is received as a reflection from another building, water body or land feature, according to the modulation; how much and to what extent will vary depending on the spectrum bands being used. In general NLOS ranging in the 2.5 GHz band mostly falls in the interval 6–8 km.

Moreover, some potentialities of WiMAX are open to the following fields of application:

- connection among Wi-Fi hotspots, and between hotspot and Internet, where *HotSpots* are the public access points, as airports, hotels and universities;
- alternative to the xDSL (*eXtensible Distance Learning System*) technology;
- high-speed services and connection for transmission from mobile devices (portables, PDA).

xDSL is a Web-based Learning Management System, which can be used to create entire quizzes online, via the Internet.

5.4.7 Search Engine

Search Engines (SE) are automatic systems able to perform searches over Web pages, so as to respond to simple user queries. Search engines support indexes of the Web pages, classifying them on the basis of mathematical algorithms defining the degree of importance of the page based on search keys.

The Search Engines operate on the Web at global, regional and local scale. The most powerful worldwide SE (Google, Live, Yahoo!, Ask, Baidu) are very few and private and, generally, the others derive their knowledge from the former. There is also some open source SE as OpenFTS, HTdig, Nutch and Egothor.

The searching capability of SE is based on three phases:

- analysis of Web pages, using dedicated crawlers;
- content cataloguing and indexing;
- computation of ranked answers to end user queries.

In the analysis phase, SE use programs called crawlers, or spiders, which automatically visit Web pages, which are uniquely identified through their *Uniform Resource Identifier* (URI) and then extract, collect and insert in databases the sensitive information, organized as metadata. The URI, also called *Web address*, is a code that univocally identifies a file, a document, an image, a service, a Web mail address.

Content classification follows the analysis and consists in inserting the pages' URI in the Search Engine indexes.

Web sites must be listed in order of relevance to answer the end user's query. The relevance of a Web site is established by searching the keyword defined by the end user. SE use *Information Retrieval* (IR) algorithms for the pages' classification using several criteria, such as

- how many times the keywords are mentioned;
- where the keywords are in the document;
- number of times the Web site is opened;
- number of links to the document.

The refinement of the Web research is based on Boolean operators (see Chapter 9). Complex searches include the choice of a document's language, the presence or the absence of phrases, date of the upgrading of the document, file format, etc.

The *Information Retrieval* (IR) is the combination of techniques used to recover in a precise way the information in an electronic format. Information means in this case documents, metadata, files, inside databases or in the WWW. Advanced aspects of IR include cognitive psychology, ontology exploitation, human behaviour with respect the information, information architecture, linguistics, semiotics. The most recent innovations in the development of algorithms and of IR systems are based on the semantic analysis of terms and on the use of semantic networks, the word meaning and the ontologies, whose role and importance are increasing.

5.4.8 Groupware

Groupware denotes software systems whose content can be manipulated by a group of users. Examples are Web sites, or collection of hypertextual documents, that can be easily modified by the users and in which the development of its content occurs in collaboration among the community that can access it, like in a *forum*. The documents are written with simple markup languages using Web browsers. The informatics branch that includes support to groupware products is the *Computer-Supported Cooperative Work* (CSCW). The term groupware assumes in the scientific literature a broader meaning, which does not necessarily include software instruments, as with the traditional teleconference.

An extension of groupware is *collaborative media*, software that allows several concurrent users to create and manage information in a Web site. Collaborative media models include *wiki* models. By method used, they are divided into:

- Web-based collaborative tools
- software collaborative tools

By area served they are divided into:

- knowledge management
- knowledge creation
- information sharing
- collaborative project management.

The term *wiki,* deriving from a Hawaiian word, *wikiwiki*, which means rapid, quick, defines the groupware used for the realization of the Web site. The first wiki Web site, *Portland Pattern Repository* (PPR), was founded by Ward Cunningham (1995) and dedicated to *People, Projects & Patterns*; it is an informal history of programming ideas. The content modification in a *wiki* is free and open and registered

in chronological succession. The scope is in sharing, interchanging, storing and optimizing the knowledge in a collaborative way.

Wiki are considered the basis on which to build a public and private stock of knowledge, and wiki technology supports the largest electronic encyclopaedias. The fields of application are several and include

- documentation
- encyclopaedias
- enterprise knowledge bases
- community wikis
- personal wikis.

5.4.9 Web 2.0 or Internet 2.0

Internet 2.0 is an evolution status, a new vision, of Internet and in particular of the *World Wide Web (WWW)*. The main aspect of Web 2.0 is that Web content is provided by users. Therefore, Web 2.0 requires interfaces easy and fast to be used, similar to the traditional software stand-alone that the users normally load on their computers. It is not specific proprietary software, but an integration of approaches for a new and innovative way to use the Web.

Web 2.0 refers to the technologies that permit the data to be independent from the producers, person or Web site. The information can be subdivided into units that move free from one Web site to others, often in ways independent from the original idea of the producer. This paradigm of the Web 2.0 allows the users to get information from different Web sites at the same time and share it on their own Web sites for new purposes.

This approach is not misappropriation of others' work for personal profit, but open source product that generates the information sharing contributing in data diffusion. New information built from the previous ones and new job opportunities are stimulated.

Web 2.0 conserves the original identity of the data, the metadata, with the possibility of modification and re-mixing by anybody for a defined purpose. Assuring data identity, the Web can be seen not only as an agglomeration but as a network of Web sites able to interact and to collectively elaborate information.

In his initial brainstorming, O'Reilly (2005) formulated the sense of Web 2.0 by the following example:

An important ingredient for Web 2.0 is *Ajax* (*Asynchronous JavaScript and XML*), a group of inter-related Web development techniques used for creating interactive Web applications on the client. Throughout the Web sites based on Ajax, the users can interact with information in each page as they are using an application, overcoming the principle of the Web as sequential navigation in static pages.

Another ingredient is RSS (*Really Simple Syndication)*, a Web content syndication format. It is a dialect of XML. The RSS allows the users to obtain automatic

Web 1.0		Web 2.0
Traditional encyclopaedias	–>	wiki encyclopaedias
Personal Web sites	–>	Blogging
Domain name speculation	–>	Search engine optimization
Page views	–>	Cost per click
Screen scraping	–>	Web services
Publishing	–>	Participation
Content management systems	–>	wikis
Directories (taxonomy)	–>	Tagging (folksonomy)
Stickiness	–>	Syndication

upgrading when a Web site changes. Through the RSS, the Web 2.0 very quickly searches, filters and re-mixes documents and images in new boxes of information.

5.4.10 Blog

The term blog *is a contraction of* Web-log*, track in the Web. Blog is a W*eb diary.

Through the blog the possibility to publish documents on the Internet becomes an opportunity for everybody, the *blogger*, and not only a monopoly of a few. The blog has some similarity with the wiki, in updating management, encouraging readers' comments and stimulating the formation of new communities.

The blog structure is based on a program that permits, when an Internet connection is available, the automatic appointing of a Webpage, without the knowledge of the HTML language, customizing with the graphical features, the template and the Web diary.

The author can publish data and information in complete autonomy. Each document is linked to a *thread*, in which each reader can write their comments and leave messages. The documents are numbered and can be identified by a permalink. The section containing links to other blogs is defined *blogroll*. The ensemble of several blogs is called a *blogsphere*.

5.5 Evolution of Hardware and Software in Geomatics

Geographical Information System is constituted by alpha-numerical data and image data, derived by satellite, airplane and/or ground acquisitions.

The increasing launch of optical and radar satellites with different resolution, the availability of digital multispectral and hyperspectral aerophotogrammetric cameras, satellite positioning systems and aircraft Laser Scanning Systems, opens a wide variety of possible dynamical territorial applications. These potentialities get nearer to the increasing need for users to update and almost real-time spatial information. The specialization induces the operators to realize ready-to-use products and the operator to access the necessary territorial data in a focused way. Resuming, the evolution derives from the following considerations:

- aircraft and satellite images and data acquisition are getting more and more diffused;
- spatial data must be made available and converted into information that can be acquired and understood by the user;
- the need for dynamic territorial analysis, in almost real time, is often of priority importance.

Some aspects are of basic importance, such as communication among the users, the release of high-quality products, data standardization, data share.

All over the world, the number of digital geographical database multiple archives is continuously increasing. To inspect the archives and to understand which one has data of interest and how it can be accessed is a quite difficult practice even for the specialists. Systems for an innovative database organization (*Spatial Data Infrastructure*, SDI) and format standardization (*Open Geospatial Consortium,* OGC) have been developed (see Chapter 9).

Thanks to the use of computer networks, the localization and access to databases has improved. The online services, including WebGIS, through the World Wide Web (WWW), are exponentially increasing. More details about these aspects are reported in Chapter 9.

5.5.1 Technology Evolution in Remote Sensed Data

In the 1960s, when the first digital images started to be processed, there were no low cost and easy to use processors. At that time technology was perforated boards and seven-trace CCT magnetic tapes; black-and-white printers and videos had very low resolution. In these conditions, even visualization of parts of a black-and-white image implied the activation of complex and time-consuming procedures. Computer architecture implied the development of sophisticated programs. The size of the windows to be analysed did not exceed matrices of 512 rows $\times$ 512 columns. In the 1970s, very few laboratories in the world had autonomous image processing stations provided with programs developed by them. The processed image restitution was another basic problem: the photographic acquisitions from the video worsened the results in terms of both colour quality and distortions. Progress in these terms has been reached by the realization of very expensive micro-densitometers which since the 1970s allowed 12.5 μm reading and writing resolutions. It was already possible to digitize an analogical image with 80 pixel resolution per millimetre corresponding to 2000 dpi (dots per inch = 2.54 cm).

Nowadays the computation power of a normal desktop personal computer is higher than that of a processor of the 1970s that was as big as a room. Moreover processing systems have become much more refined, have a large market and are getting much easier to use. Nevertheless the quantity of data to manage and pro-

cess are in the order of Giga–Tera–Petabytes, and visualization and processing are requested in reasonably short times (Box 5.2).

Box 5.2 Evolution of the airborne and satellite remote sensing systems used in geomatics

Year	Detection system
1858	Gaspar Felix Tournachon, aka Nadar, captured the first aerial view of Paris, from a balloon at an altitude of 400 m
1859	Starting of the photogrammetry; Aime Laussedat presents at the Commission of the Academy of Science in Paris a report concerning the possibility to calculate coordinates of an object based on the spatial intersection of perspective rays referred to a couple of photograms (stereo-pairs) In 1849 he was the first person to use terrestrial photographs for topographic mapping
1893	The word *photogrammetry* was first mentioned in a publication by Albrecht Meydenbauer
1964	Activation of the satellite positioning system TRANSIT, progenitor of the Global Positioning System (GPS), USA
1968	Hemphill describes the use of the laser in the aerial survey
1972	Launch of the first satellite ERTS-1 (Landsat-1) for the study of Earth's resources, with acquisition capability in four spectral bands (MSS)
1973	Beginning of the operativeness of the global positioning program GPS for military scope
1982	Beginning of GLONASS, Russian global positioning system, operativeness
1982	Launch of the Landsat-4 Thematic Mapper (TM) with improved geometric (30 m) and spectral (7 more selective bands) resolutions
1986	SPOT- 1 is in orbit, with electronic acquisition system, stereo capability, panchromatic band with 10 m geometric resolution
1991	The European Space Agency launches the satellite ERS-1, with an active system SAR (Synthetic Aperture Radar)
1990s	Beginning of the operativeness of the hyperspectral aerial sensors
1995	Launch of Radarsat-1, first commerical active Earth observation satellite
1996	NASA's *Mission To Planet Earth* (MTPE) undertook the first of a planned series of biennial reviews to assess its end-to-end observations/research/data management programme, including *Earth Observation System* (EOS)
1997	Beginning of the commercial satellite launches with high geometric resolution (<3 m pan) and with flexibility in the stereoscopic acquisition
1999	Start of the ESA programme *Living Planet*, including the *Global Monitoring for Environment and Security* (GMES), renamed Copernikus in September 2008

2000	Presentation of the first aerial photogrammetric digital cameras (ISPRS 2000, Amsterdam)
2002	Launch of the satellite ENVISAT (ESA), carrying optical and radar sensors, for global and regional studies of Earth's resources and of the atmosphere
2003	Start of the COSMO/SkyMed system for a constellation of optical and radar satellites for Earth Observation (Italian Space Agency)
2003	Approval of the Galileo project for a European global positioning system
2008	Launch of GeoEye-1, satellite with the highest resolution of any commerical imaging system of 0.41 m in the panchromatic mode

5.5.2 *Configuration of a Geomatics System*

To define the configuration of a system that should be adopted to improve professional quality, the following questions have to be asked:

- What are the purposes and what kind of product is intended to be realized?
- What are the current and future human and financial resources available?
- What are the input territorial data available, their quality and homogeneity?
- If it is intended to realize a GIS, what are the possibilities for future acquisition of data and their update?
- What software systems are intended to be used?
- Are software systems used interoperable?
- Are they compatible with the already available data or databases that are intended to be used?
- How much is accounted for investing in training and information to follow the quick evolution both of the techniques and of the technologies?

When buying an informatics product, everybody should know that the innovative product can already be old tomorrow both from a purely technological point of view and in terms of functioning logic. In general the focus should be on the most recent products with attention to the possible evolutions of each component, their compatibility and future updating.

5.5.2.1 Working Station

Processors performances do not depend only on the CPU (*Central Processing Unit*) but also on the harmonic balance of all the internal sub-systems, like hard disk, memory, cache, video board, mother board architecture and operating system, as well as from the external input.

One or more internal components, software and hardware, can be of limitation to the others, slowing down the system.

Although the increasing overall capacities and the simultaneous decrease of costs, an incorrect configuration can produce reduced potentialities and limited performances.

Another basic element is the video. Higher are performances, memory board, colours resolution, etc., better are the visualized results.

The connection of the working station to the Internet and Intranet networks, in order to have the possibility to access the services of data transfer via software (e.g. via FTP) and to access information and archives in distributed databases (e.g. access to shared databases, research of the available images and of their metadata, geographic coordinates included, identifying codes) represents the quickest way to interactively exchange information and to acquire data. If the intention is to work with many units, the internal networking connection facilitates the information interchange reducing the time for the planned operations.

Figure 5.18 reports a configuration scheme of a working station for digital image processing and GIS development. The central unit is a personal computer.

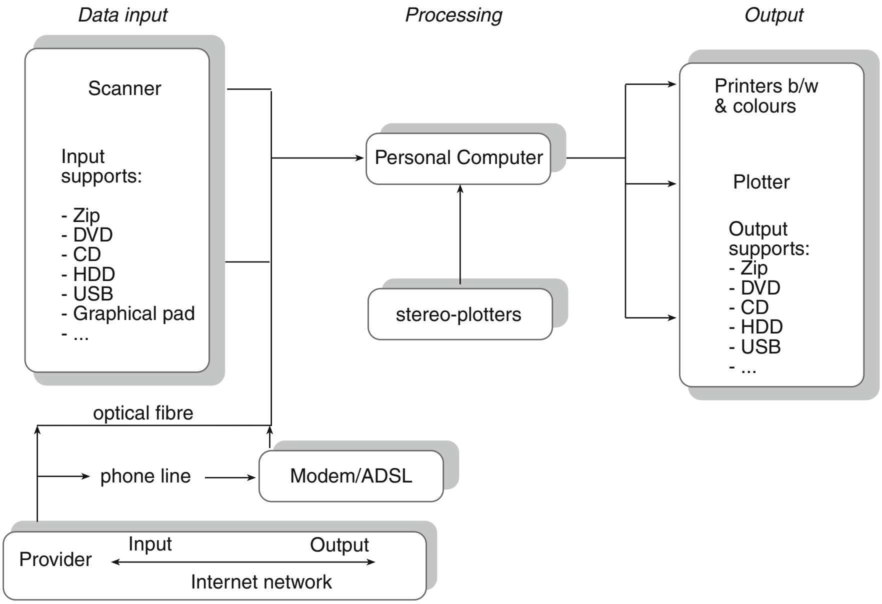


Fig. 5.18 Geomatics working place configuration for personal computer with some input/output devices

5.5.2.2 Consideration About Software

To be efficient at its best, a professional working station has to be equipped just with programs necessary to the established purposes, to avoid occupying it with program packages that usually only have the function to subtract computation power.

The choice of the most suitable package of programs for image processing and/or GIS, as has been said, is a function of the use needs. In general it is desirable to have the most common functions for 2D, 3D and 4D image processing (the fourth dimension is the time), for data management in a GIS as well as for vector data complete management. In particular, special attention has to be paid to the input data readable formats and to the output formats, in order to be compatible with other systems.

As regards the different types of GIS software available on the market, three levels of complexity and performance can be distinguished:

- *the most economic level*, with program packages, usually developed by Universities or Research and Training Institutes, that have the most common basic functions and in some cases even with specific capacities for a certain application in environmental topics, and with vector and/or raster GIS functions. They are very useful at both university level and specific professional level training;
- *derived from the higher level software*, grouping one or more specific potentialities and for focused applications;
- *more advanced and complete level* for a wider possibility of applications in different domains, generally developed in excellence research fields for highly specialized working sectors.

5.6 Summary

This chapter introduces the domain of computer science, contributing to understand the basic principles of the most techniques and instruments used in the field of geomatics.

The dynamics of computer science by presenting topics in a historical perspective in which past developments, the current state of the art and directions of research are discussed. The coverage of topics such as programming languages, operating systems, algorithms, software engineering, networking, database design, artificial intelligence and machine architecture result in an overview of the field of computer science presented in brief.

The fundamentals of computer science are based on the concept of computers for general applicability developed by John Von Neumann in 1945. The conceptual elements are the processing unit, the central memory, the secondary or mass memory, the peripherals and the bus system. The machines stored instructions as binary values (creating the stored-program concept) and executed instructions sequentially one at a time and processed them. The main components are the hardware, the physical components of the system, and the software, programs executed by the system. Processing operations are performed by an algorithm, which describes a sequence of precisely defined steps leading to the realization of a task.

Processors have computation flexibility and can be classified, in relation with their processing capabilities, in personal organizers, PDAs, notebook, workstation and mainframes.

The machine language is a programming language operating on the computer physical components. High-level programming languages have been created to simplify the user access operating on symbolic components of the machine. Computer basic software is the operating system (OS) representing the set of programs making the processor operative and usable; DOS, MacOS, Windows, UNIX and Linux are briefly introduced.

Databases are structured collections of organized data that allow inserting, modifying, cancelling, organizing and restoring of data. Databases can be organized in different models as hierarchical, network, relational and object oriented by specific software called DataBase Management System (DBMS). Specific software for relational and object-oriented models are the Relational Database Management System (RDBMS) and the Object-Relational DataBase (ORD).

The necessity to transfer data and information at a distance, diffused and accessible to everybody, stimulated the conditions to develop computer networks. There are different transmission modes as Unshielded Twisted-Pair (UTP), co-axial cable, optical fibre and wireless, through EM waves. Data transmission can be extended worldwide, Wide Area Networks (WANs), or for local purposes, Local Area Networks (LANs). Communication protocols are necessary to define conventional rules for data representation, signalling, authentication and error detection.

The most common network infrastructure is the Internet, the most important geographic network worldwide, and the World Wide Web (WWW) is the widest set of informatics objects available on the Internet. The virtual links are supported by high-level communication protocols which allow transferring hyper-objects through the network, as the HyperText Markup Language (HTML).

An Intranet is a local, private network realized by the application of Internet high-level standards and protocols.

The success of computer diffusion depends also on wireless technology, WiMAX (Worldwide Interoperability for Microwave Access), permitting access to broadband telecommunication networks without wires.

Other utilities of increasing importance are the Search Engines (SE), the Groupware, the Web2.0 or Internet 2.0 and the Blog.

The scope of this book is to introduce the basic elements in geomatics. One fundamental step is the configuration of a working station for geomatics development and application. Some elementary notices are described in a short introduction concerning the evolution of the systems in particular for Earth Observation image processing techniques and Geographical Information System.

Further Reading

Brookshear J.G., 2005, Computer Science. An Overview, 8th ed. Addison-Wesley, Reading, MA, p. 576, ISBN-10: 0321247264.

de Caluwe R., de Tré G., Bordogna G., 2004, Spatio-Temporal Databases, Springer.

Heims S.J., 1980, John von Neumann and Norbert Wiener: From Mathematics to the Technologies of Life and Death. The MIT Press, Cambridge, Massachusetts Berlin.

Musser J., O'Reilly T., 2006, Web 2.0 Principles and Best Practices, O'Reilly Radar Report, p. 101, ISBN 0-596-52769-1.
O'Regan G., 2008, A Brief History of Computing. Berlin Springer, p. 252, ISBN: 978-1-84800-083-4.
Renaud P.E., 1996, Introduction to Client/Server Systems, A practical Guide for Systems Professionals. John Wiley & Sons, New York.

Bibliography

Aspray W., 1990, John von Neumann and the Origins of Modern Computing. MIT Press, Cambridge, Massachusetts.
Codd E.F., 1970, A relational model of data for large shared data banks. Communications of the ACM, 13(6): 377–387.
Elmasri R., Navate S.B., 1999, Fundamentals of Database Systems, Ed. Benjamin/Cummings Redwood City.
Gallippi A., 2001, Dictionary of Informatics and Multimediality/Dizionario di Informatica e Multimedialità, English/Italian, Ed. TecnicheNuove.
Melton J., Simon A.R., 1993, Understanding the New SQL: A Complete Guide. Morgan Kaufmann, San Francisco, CA.
O'Reilly T., 2005, What Is Web 2.0, Design Patterns and Business Models for the Next Generation of Software.
Rigaux F., Scholl M., Voisard A., 2002, Spatial Databases, with Application to GIS. Morgan Kaufmann Publishers, San Francisco, CA.

Chapter 6
Acquisition Systems

The technique of image acquisition of the terrestrial surface can be applied using tools on board different platforms.

Three kinds of platforms are generally proposed: (Fig. 6.1):

- *satellite platforms*: the Earth Observation (EO) from orbital distances is the principal characteristic of remote sensing from space, both from shuttles with crew

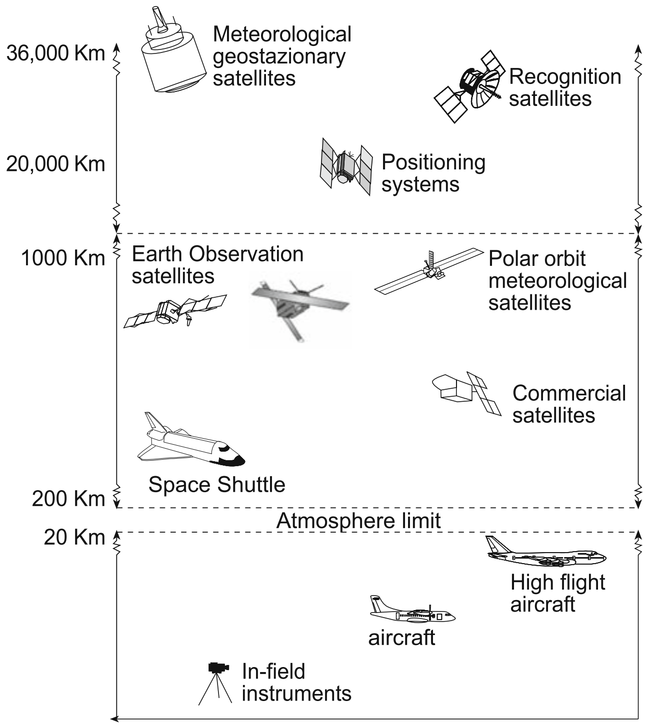


Fig. 6.1 Platforms for the Earth Observation and satellite global positioning systems with reference to the operative orbital altitude

M.A. Gomarasca, *Basics of Geomatics*, DOI 10.1007/978-1-4020-9014-1_6,

and automatic satellites. The altitudes range from 200 km, for space shuttle and shuttles with crew, to 35,800 km, for geostationary orbit of meteorological satellites. For what concerns satellites, and particularly those used to observe and to study the terrestrial resources, the altitude varies from 450 to 900 km, in sun-synchronous, near-polar and circular orbits (see Plate 6.2);

- *aerial platforms*: consist of bounded or free balloons (which carry the instrumentation in a special box), airplanes, helicopters, dirigible, etc. The most widely used is the airplane which can be employed from low to medium heights (from 300 to 6000 m) for local or for small-area observations from a higher altitude (up to 20,000 m), in order to obtain acquisitions over wider areas;
- *ground platforms*: the methods for remote observation which use instrumentation based on the terrestrial surface. An example is a vehicle with an overhead jib carrying the instrument for the survey; in this way, a vertical acquisition can be obtained, at the nadir point, up to 15 m height; its limit consists in the fact that the vehicle can operate only on roads and paths. Alternatively, we can use detachable and transferable metallic towers. The observation from these platforms is useful to collect data such as the spectral responses of the surfaces, to set and to calibrate aerial and satellite images, to know through continue observation the behaviour of the temporal spectral reflectivity and to correct atmospheric effects.

6.1 Imagery Generation

Analogical photographic images, recording the electromagnetic radiation in the visible interval of the spectrum, are the first source of data for terrestrial surface observation. Through image interpretation, mainly based on soil texture analysis, it is possible to discriminate classes of land cover.

With digital techniques, the analysis is mostly based on pixel values for each spectral band. Analogical and digital acquisition processes are conceptually identical, completely different from the components that lead to the formation of photographic or digital images (Fig. 6.2).

The development of instruments for recording multispectral signals allowed the acquisition of quantitative information concerning the objects on the Earth's surface (Fig. 6.3).

The automatic satellites in continuous orbits facilitated the acquisition and transformation of multispectral signals into digital format (Fig. 6.4).

A sensor detects an electromagnetic signal in analogical format according to a variable function. Since the speed of scanning is constant and known, such function can be graphically represented, reporting in x-coordinate both time and space and in y-coordinate the representation of electric tension, since the majority of the physical units can be translated in electric units.

The corresponding tensions are measured at regular intervals by sampling the analogical signal (Fig. 6.5). The different tension values, read in correspondence of regular intervals, are measured by an electronic voltmeter and converted into numerical value by an analogical/digital converter (see Plate 4.1).

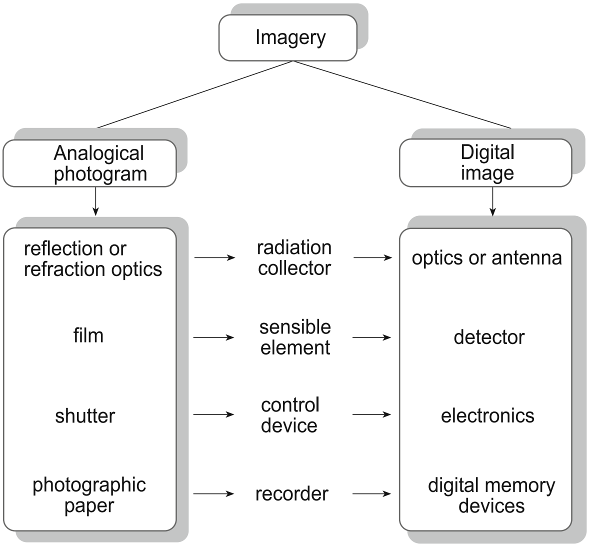


Fig. 6.2 Analogical (photographic) and digital (scan) processes to obtain images; the two approaches have common roots

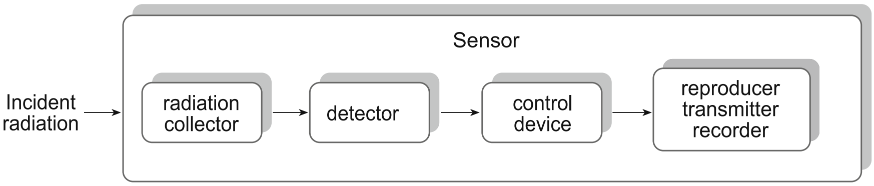


Fig. 6.3 Conceptual design of a sensor, both analogical and digital: The sensor collects the information conveyed by the electromagnetic energy. The detector is the component of the sensor able to transform the incident electromagnetic radiation in a corresponding electrical signal

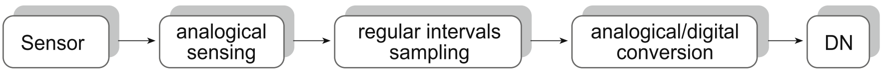


Fig. 6.4 From the radiation physical analogical data collection to the digital binary conversion

In the sampling operation, information between subsequent measurements is generally lost. Shannon (1949) demonstrated that if the sampling frequency is double of the maximum Fourier frequency of the signal, the information is not lost.

The analogical signals are transformed into numerical values in function of the established sampling frequency. The more frequent are the samplings, the more exact is the numerical description of the original analogical signal, though the amount of data to be managed increases.

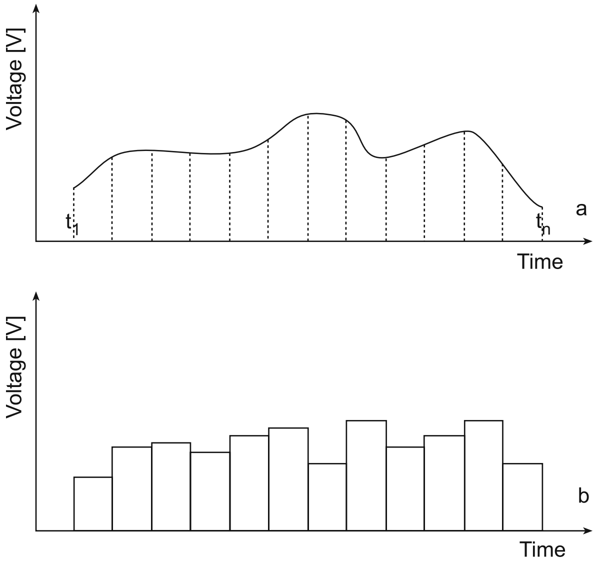


Fig. 6.5 (**a**) The analogical signal is continuous in the time and linear with respect to the observed phenomenon. The analogical signal is uniformly sampled at regular intervals from t_1 to t_2; (**b**) discrete sampling of the analogical signal. Thickening the sampling improves the radiometric resolution, increasing the volume of the data

The signal changes from continuous to discrete.

This discretization defines the *radiometric depth* of the pixel or *radiometric resolution*, expressed also as other synonyms:

- bit resolution;
- chromatic or tonal scale;
- dynamic range or field;
- dynamic elements;
- intervals of Digital Number (DN).

The radiometric resolution indicates the number of bit associated to each pixel of the image and determines the chromatic or tonal scale (Fig. 6.6).

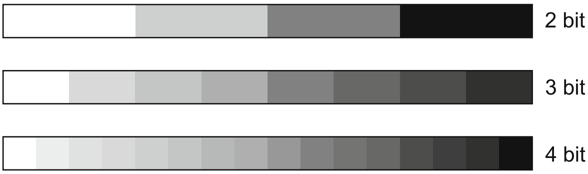


Fig. 6.6 Chromatic range, or radiometric resolution, in grey scale of 2-, 3- and 4-bit image

The elementary unit of information, called bit (binary digit), is composed of two digits 0 and 1, corresponding to the possible states:

1: on

0: off

A numerical datum can be represented by an ordinate series of bit. The possible numerical combinations in 2-bit image are four: 00, 01, 10, 11. For every further bit, the number of combinations is raised to the square (2^k, where k is the number of bit); in 4-bit image, the possible combinations rise to 16 (2^4), in 8-bit combination to 256 (2^8), and so on.

If each colour, blue, green, red, is represented by 256 different tonal levels, the images can be represented with more than 16 million colours ($256 \times 256 \times 256$).

The analogical–digital conversion is necessary when the transmission of analogical signals from satellite or airplane to a ground-receiving station suffers irreversible distortions of the scanning signals due to the atmospheric path. The transmission of a digital signal is less critical because the possible states 1 and 0 can be recognized and reconstructed.

6.1.1 Charge-Coupled Device (CCD) Detector

The most used sensor for the realization of different types of instruments on board platforms is the *Charge-Coupled Device Detector* (CCD), which measures and records the luminous energy using the digital conversion process.

An electronic circuit scans the matrix by lines and columns (Fig. 6.7), providing an electric output signal containing, point by point, the measure of each photodiode.

When a CCD is exposed to the light, a flux of electrons proportional to the intensity of the incident radiation is released. This flux induces a variation of the analogical electrical signal. The recording and the following digital sampling allow the

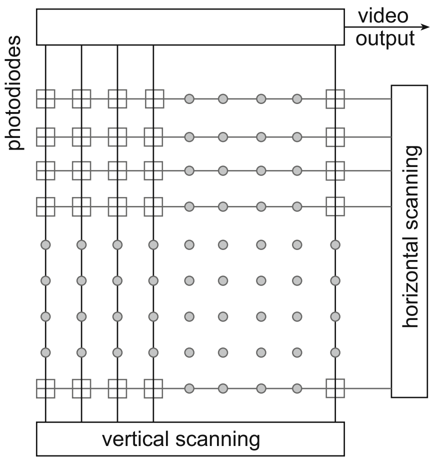


Fig. 6.7 The Charge-Coupled Device (CCD) scanning system

representation in Digital Numbers: the larger the emission of electrons, the greater the radiance, represented by a higher DN.

Some colour filters, selectors of luminous radiation, are set in front of the CCD determining the sensibility of the sensor to a given wavelength (λ).

Compared to photography, the image through digital acquisition more rapidly obtains better chromatic accuracy. This permits the elimination of the undesired dominants often present in the film.

6.1.2 Acquisition Geometry

Passive scanning systems are used for the acquisition of ordered series of surface stripes, normal to the satellite motion, in one or more bands of the electromagnetic spectrum.

Passive systems are characterized by three ways of data acquisition, according to the mechanisms used:

- *central perspective acquisition*: the whole scene is visualized from a point of view located in the centre of the image. This system, which increases geometrical distortion on reaching the borders of the scene, is adopted in the *Return Beam Vidicon* (RBV) sensor, multispectral camera on board the first series of Landsat and of some geostationary satellites;
- *rotating mirror optical–mechanical system*: it collects data through a sensor connected to a mirror moving in a direction perpendicular to the platform motion; at the end of each oscillation, a swath with a number of pixels functional to the sensor resolution is obtained. The main limitation of this acquisition mode is the little time available at the detector to collect data, for which it is necessary to continuously keep the signal-instrument noise ratio controlled. NOAA-AVHRR, Landsat MSS, TM and ETM+ sensors carry this mechanical scanning device;
- *electronic elements in series along an array*: *whiskbroom* and *pushbroom* sensors. This solution improves the signal-to-noise ratio thanks to the short acquisition time. Many satellite and photogrammetric acquisition systems, photographic digital cameras as well as instruments for ground sensing use the CCD technology. The acquisition geometry depends on the physical disposition of the detectors. The acquisition geometries can be defined as follows:
  - *central matrix*;
  - *punctual scan* (*whiskbroom*) non-central;
  - *linear scan* (*pushbroom*) non-central.

The digital images produced with punctual (*whiskbroom*) or linear (*pushbroom*) CCD scanning systems exclude the instantaneous acquisition of the whole scene (simultaneity of the acquisition).

The *whiskbroom* (single cell) and the *pushbroom* (line) geometries are more complex than in the classical central perspective. In fact, there is a lack of one unique

acquisition centre for the single image, which can be seen as a mosaic of consecutive sub-images each one having their own perspection centre.

In such a scene, there is not a unique central perspective, but many. Phenomena of variation of the six external orientation parameters of each single central perspective must be modelled, considering, with satellite acquisition, the orbits and Keplerian orbital motion perturbations (see Box 7.1).

6.1.2.1 Central Matrix Acquisition Geometry

In the *matrix sensor,* the detectors are structured in a grid according to square or rectangular shapes. The sensor acquires the image in only one exposure. More exposures could be necessary for acquiring multispectral data.

The operational mode is similar to the traditional photogrammetric instruments because the scene is acquired as a whole and not point by point or line by line (Fig. 6.8).

6.1.2.2 Punctual Scan Sensor (Whiskbroom)

The punctual scanning system counts the use of a single detector (CCD).

The *whiskbroom* system is an across-track scanner. It uses rotating mirrors to scan from side to side perpendicular to the direction of the sensor platform. The width of the sweep is referred to as the sensor swath. The rotating mirrors redirect the reflected light to a single or a few detectors. Whiskbroom scanners with their moving mirrors tend to be large and complex to build. The moving mirrors create spatial distortions that must be corrected by image pre-processing (see Chapter 8). The fewer number of detectors to keep calibrated is an advantage of whiskbroom scanners, as compared to other types of sensors (Fig. 6.9).

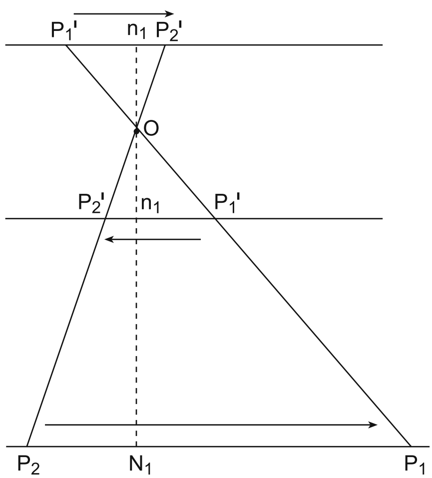


Fig. 6.8 The central perspective concept

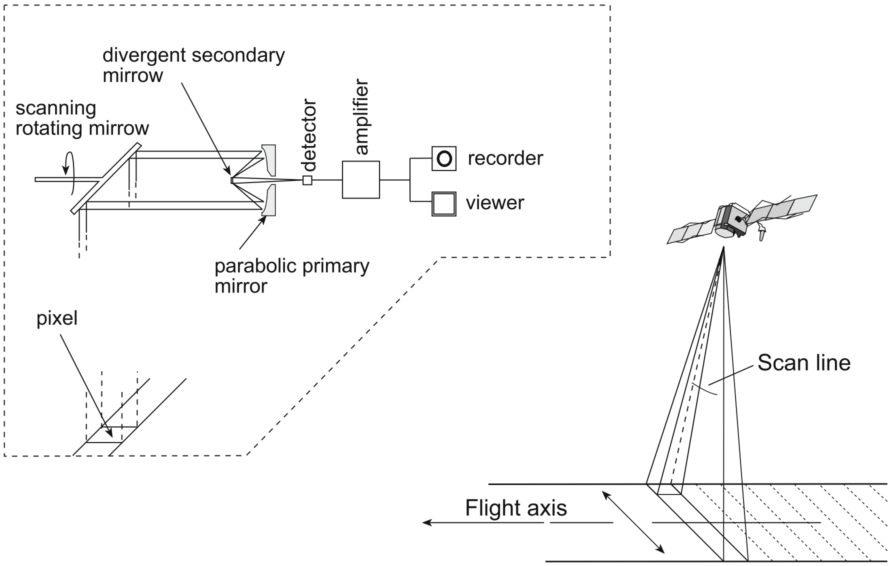


Fig. 6.9 Data collection and digital restitution of a point by point optical – mechanical scanning device (*whiskbroom*). The Landsat satellites and the airborne hyperspectral MIVIS use this technology

Satellite sensors of the first generations and some hyperspectral airborne systems belong to this typology, while it is less used in photogrammetry.

6.1.2.3 Linear Scan Sensor (Pushbroom)

The linear sensor is made by an array of CCD detectors. The scanning of the object is carried out line by line. To allow the complete RGB chromatic acquisition in one step, a three linear array CCD is used, in which each detector array has respectively red (R), green (G) and blue (B) filters (Fig. 6.10).

The possible configurations of such systems are

- *pushbroom*;
- *panoramic*.

In both cases, the sensor is a rigid support of detectors able to simultaneously acquire many cells of the same scanning line.

Pushbroom geometry: the sensor is usually perpendicularly oriented to the direction of the motion of the platform. The linear solution allows the elimination of mobile mechanical devices in the scanning phase (Fig. 6.11).

The central perspective is maintained line by line (along the direction perpendicular to the motion of the platform) obtaining an orthographic image in the direction of the flight. Each line produces an independent image characterized by its own centre of acquisition and attitude. SPOT, IRS and the generation of commercial satellites are structured with these systems.

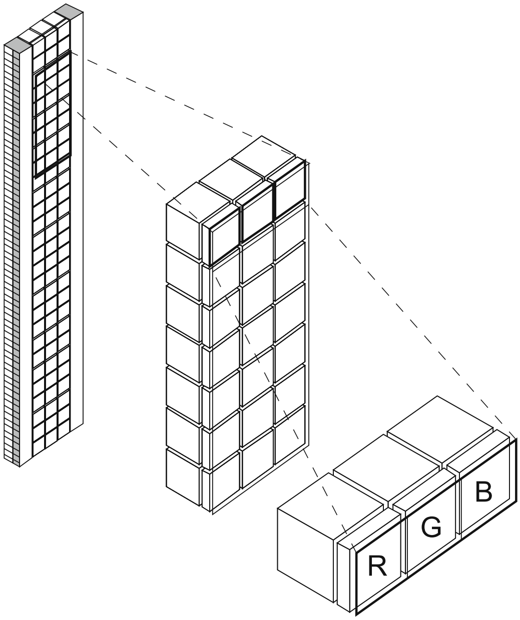


Fig. 6.10 Three-linear digital sensor

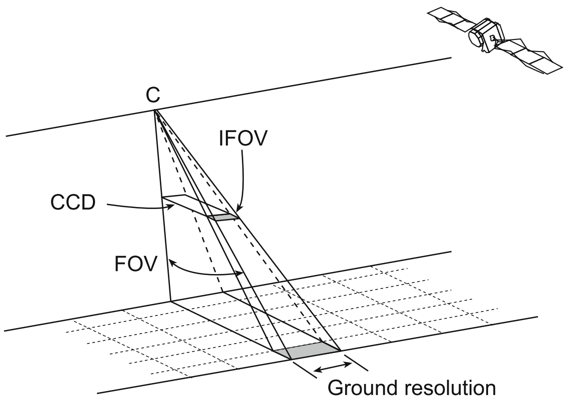


Fig. 6.11 Acquisition geometry of an electronic linear scan system (pushbroom). The linear array is constituted by several CCDs (Charge-Coupled Devices) which work simultaneously along the scanning line

Panoramic geometry: the linear sensor is perpendicular to the direction of the flight, and it scans perpendicular to the ground. Such systems allow the coverage of large areas with high resolutions avoiding complex lens systems.

The scanning of each line must be completed in the time necessary at the platform to cross the field of view in the direction of the flight. Stereoscopic acquisition is possible even if difficult.

The geometry of the output image is very complex (Fig. 6.12):

- the image scale changes because of the high variability of the distance sensor-object between nadir and horizon positions;

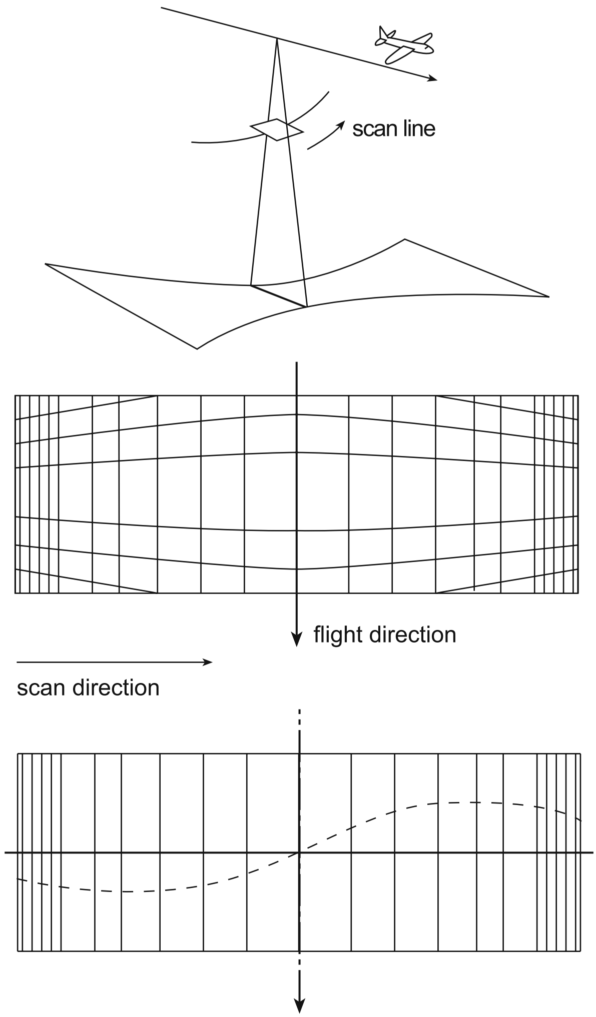


Fig. 6.12 Panoramic linear scanning system; a: operative logic; b: distortion due to the transversal panoramic scan of the scene; c: distortion caused by the motion of the platform on the scene

- the transversal movement of the scanning matched with the advance of the platform induces further deformations;
- motion compensation systems also contribute to the complexity of the geometry.

6.2 Instrument Resolution

Each instrument is characterized by different resolutions.

The *geometric resolution* (Fig. 6.13) is related to the ground size of the elementary cell; a remotely sensed image is made by basic elements called pixels (picture

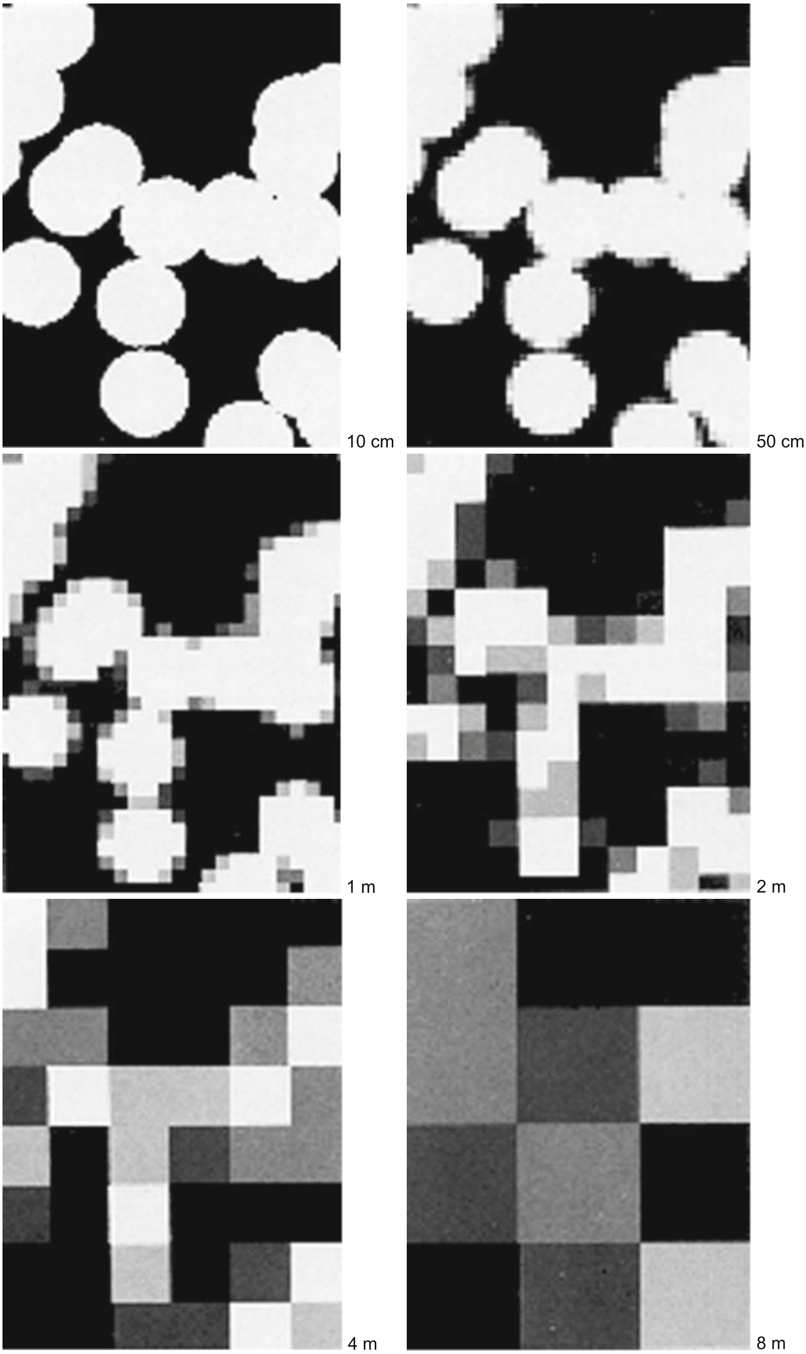


Fig. 6.13 Comparison of images of the same scene acquired with different geometric resolution

elements). Given a digital image, each of its elementary surfaces is a pixel. Each pixel is characterized by two coordinates (x, y) indicating its position in the image and by a Digital Number (see Chapter 8).

Pixel size influences the resolution of the scene details and is related to the distance of the acquisition system, the characteristics of the sensor and the type of operation. The geometric resolution can vary from centimetres up to several kilometres.

The pixel is also defined as *Instantaneous Field of View* (IFOV), while the *Field of View* (FOV) is the matrix of the image pixels (Fig. 6.14). The IFOV is the solid angle within which radiation is detected at a given instant.

The geometric resolution needs to be defined in metric terms to avoid frequent misunderstanding between operators and users when generically referring to high, medium or low resolution (Fig. 6.15). An indicative classification is summarized in Table 6.1.

The *spectral resolution,* or resolving power, is a measure of its power to resolve features in the electromagnetic spectrum. It is usually defined by:

$$R = \lambda/\Delta\lambda \tag{6.1}$$

where $\Delta\lambda$ is the smallest difference in wavelengths that can be distinguished, at a wavelength of λ.

The spectral resolution is also related to the number of spectral bands and their correlation.

Radiometric resolution, or radiometric sensitivity, refers to the number of digital levels used to express the data collected by the sensor. The *radiometric resolution* is the minimum energy able to stimulate the sensitive element in producing

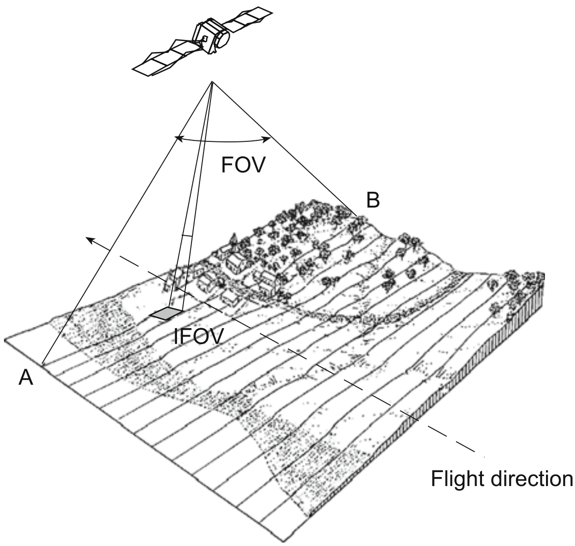


Fig. 6.14 Sensor scan swath: the angle of view is defined as Field of View (FOV). The spatial resolution, i.e. the angular aperture corresponding to a pixel at the nadir, is defined as Instantaneous Field of View (IFOV). AB: scan swath

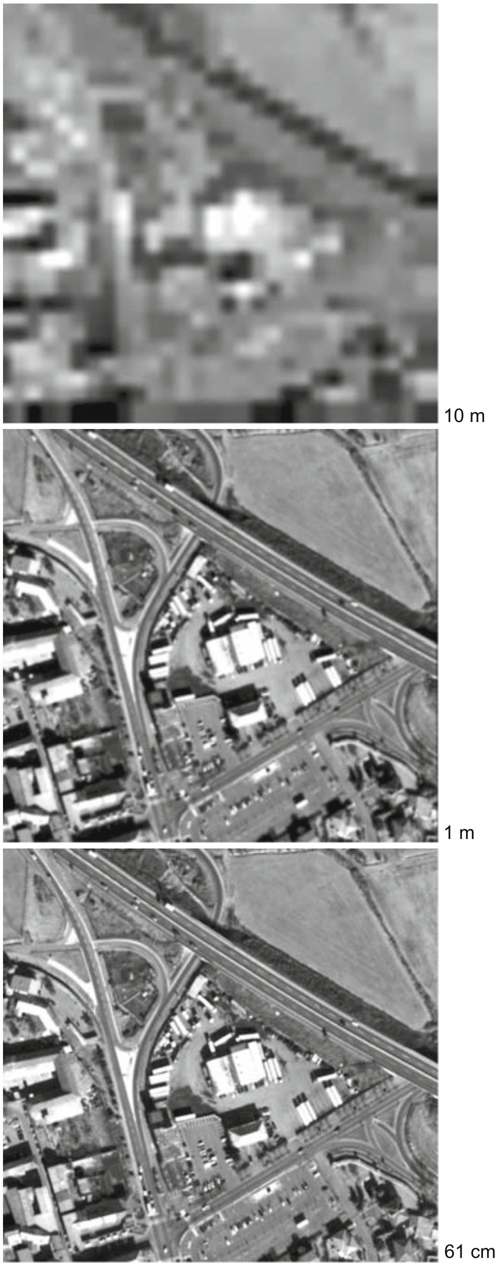


Fig. 6.15 Images of the same scene acquired with different geometric resolution: (**a**) 10 m, (**b**) 1 m, (**c**) 61 cm (highway interchange, Turin Caselle, Italy)

Table 6.1 Classes of geometric resolution for airborne and satellite acquisition systems

Class of resolution	Geometric resolution (m)	Resolution	Acronym
1	0.1–0.5	Very very high	VVH
2	0.5–1	Very high	VH
3	1–4	High	H
4	4–12	Medium	M
5	12–50	Medium low	ML
6	50–250	Low	L
7	250–1000	Very low	VL
8	>1000	Very very low	VVL

a detectable electric signal, beside the intrinsic noise, in relation to the sensor sensibility to detect the intensity of the electromagnetic signal coming from the detected objects. There is a minimum radiance interval $\triangle\lambda$ that is in a Digital Number (Fig. 6.16).

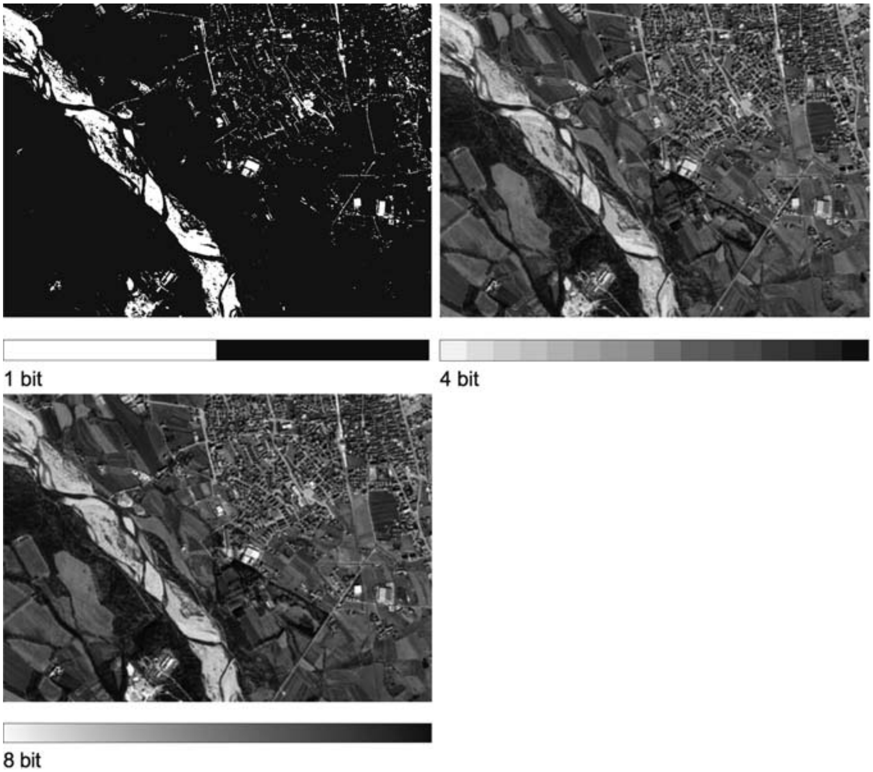


Fig. 6.16 Images of the same scene acquired with different radiometric resolution

The *temporal resolution* is the period between two following acquisitions of a same area. Figure 6.17 shows the effect produced by the different inclination of sun rays on the typology of shades generated in the same area in different seasons, reported on a series of Digital Elevation Model (DEM).

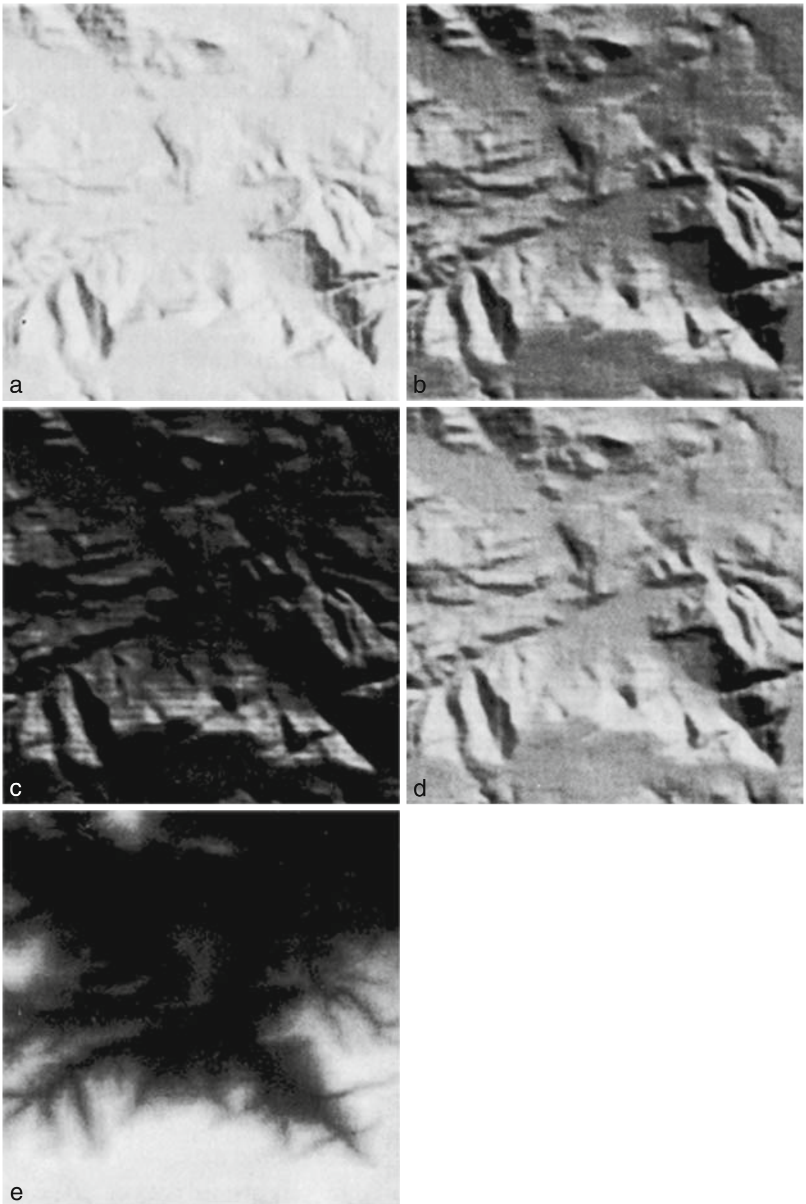


Fig. 6.17 Digital Terrain Model (DTM) of the same scene (**e**) and different reflectivity in the four (**a**, **b**, **c**, **d**) seasons at 45° latitudes

6.3 Earth Observation Satellites

Orbiting stations are planned, years in advance, trying to realize the results of scientific research and to foresee the market requests. The satellites in orbit now are the result of researches started in the 1960s, when the use of images was reserved for research laboratories, companies in the sector and a few others who invested, with high costs, in high technology and specialists.

An important element to evaluate is the time for delivering to the users; in fact some uses require information in real or near real time. This decisive element has influence on the means and methods of transferring and processing information and on the cost of structures and instrumentation.

The estimate for the future is a request of images having better geometric, spectral, radiometric and temporal resolution, with an increase of received, transmitted and processed data (gigabit per second). Specific plans derive both from the possible applications and the user's needs (Box 6.1).

Box 6.1 Land management and potentiality of the satellite imagery as support for decision making

Application	Geometric resolution (m)	Spectral region				Type of acquisition	Period
Hydrology							
Soil moisture	5–30		M		R	m	7 days
Rainfall	250–1000		M		R	m	1 day
Hydrological balance	30		M			m	1 day
Snow cover	5–120	P	M	T	R	m	7 days
Water equivalent from snow	5–120		M	T	R	m	16 days
Artificial watershed census	4–30	P	M		R		Annual
Artificial watershed monitoring	4–15	P	M		R	m	16 days
River and lake water level	4–15	P	M		R	m	1 day
Drainage network	1–30	P	M		R	m	6 months
Soil							
Soil unit	1–30	P	M			m-s	Several years
Erosion qualitative evaluation	0.5–20	P	M		R	s	3 months
Agriculture							
National inventory	5–30		M		R	m	3 months
Local inventory	0.5–15	P	M		R	m-s	15 days
Yield forecasting	5–200		M			m	7 days
Vegetation							
National inventory	5–30		M		R	m	3 months
Local inventory	0.5–15	P	M		R	m-s	3 months
Forest damage	3–120		M	T		m	7–30 days
Fire monitoring	0.5–30	P	M		R	m	Few hours
Fire inventory	0.5–30	P	M		R	m	3 months
Photogrammetry							
Map >1 :50,000	10–200	P	M				
1 :5000−1 :50,000	0.5–10	P	M				
Thematic map >1:25,000	4–200	P	M		R		

1:5000–1:25,000	0.5–15	P	M				
DEM generation	1–20	P	M		R	s	1 year
Geology/geomorphology							
Lithology	1–30	P	M			m	Several years
Structural analysis	15–200	P	M			m-s	Several years
Geomorphologic unit	1–30	P	M		R	m-s	Several years
Hydrogeological asset	0.5–20	P	M		R	s	1 month
Geothermic	10–120	P	M	T	R	m-s	6 months
Archaeology							
Paleo-environmental indices	0.5–20	P	M		R	m	several years
Heritage texture	0.5–20	P	M		R	m	several years
Pollution							
Landfill identification	0.5–15	P	M		R	m-s	6 months
Landfill census	3–15		M			m-s	3 months
Pits census	1–20	P	M		R	m-s	6 months
Coastal pollution	1–20	P	M			m	7 days
Marine pollution	15–200		M	T	R	m	1–7 days
Inland water quality	3–10		M			m	1–7 days
Risk							
Flooding prevention	1–30	P	M		R	m-s	3 months
Flooded areas monitoring	0.5–30	P	M		R	m	1 day
Flooded areas surface estimation	0.5–15	P	M		R	m	1 day
Seismic risk							
Faille identification	15–30	P	M		R	m-s	3 months
Damages assessment	15–30				R	m	Real time
Volcano monitoring							
Eruption monitoring	1–30–120	P	M	T	R	m-s	Real time
Damaged surface estimation	0.5–30	P	M		R	m-s	7–15 days
Human activity							
Urban typology	0.5–15	P	M			m-s	Annual
Population density	0.5–20	P	M			m-s	Annual
Primary road network	1–30	P	M			m-s	Annual
Secondary road network	0.5–5	P				m-s	Annual

6.3.1 History of the Space Missions

Conquering Space has always been an intriguing challenge for man throughout history. From Icarus's foolish dream, through Astolfo's mythical flights to Paradise and to the Moon by a fantastic creature, the Hippogryph, seeking the hero Orlando's sense, madly in love but unrequited by Angelica, in the *Mad Orlando* by Ludovico Ariosto, to the incredible fantasies of Jules Verne in *From the Earth to the Moon* which, in a child's mind, became the conviction that the fantastic can be realized (Fig. 6.18).

Hermann Julius Oberth (Transylvania 1894–Nuremberg 1989) planned the first prototypes of rockets, the V-2, during his long life, and took part in the development of the US space plan after the Second World War.

In 1923, opposed by the scientific world (his PhD dissertation was not accepted) he published the first book about the theory of positioning satellites in orbit around

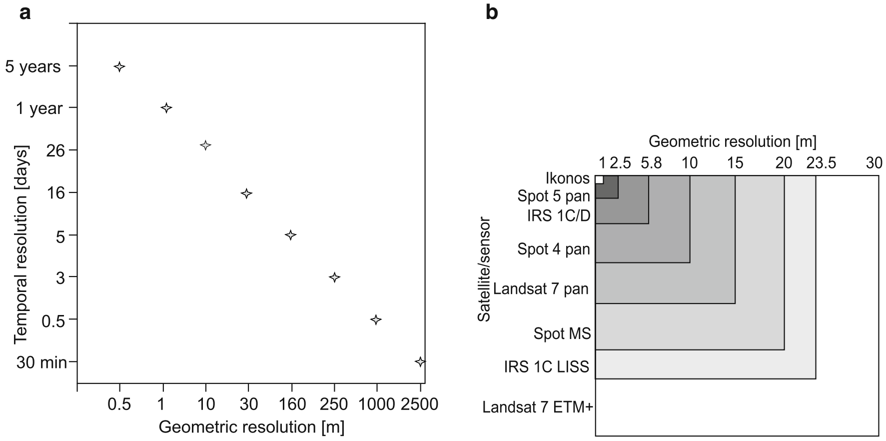


Fig. 6.18 (**a**) Relationship between temporal and geometric resolutions in some of the most used satellite acquisition systems (nadir). With along track or lateral pointing view systems, the revisiting period can be reduced generating however more geometric distortion; **(b)** the geometric resolution (pixel) of some satellite sensors. QuickBird, not indicated, has 0.61 cm side pixel, about ¼ of IKONOS' pixel size

the Earth. This was recognized as research of fundamental importance for humanity. A more complete version followed in 1929.

With the important help of his student *Wernher von Braun* (Fig. 6.19) in 1942, the first long-range rocket V-1 was tested and later in 1944 the rocket V-2, weighing 12 tons, was launched at 800 km altitude.

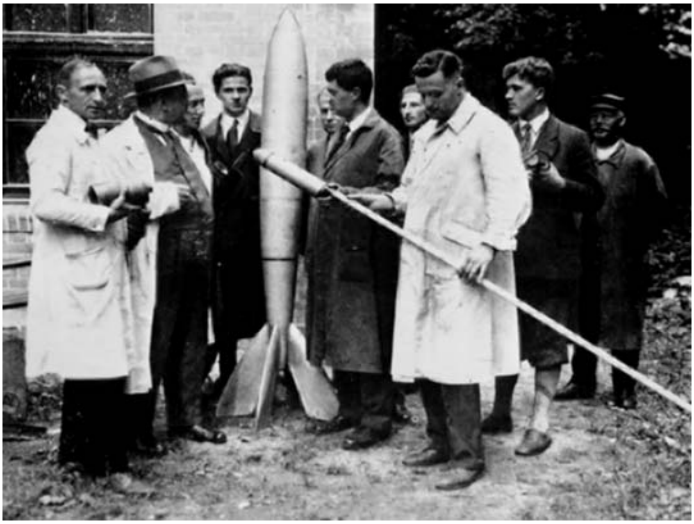


Fig. 6.19 Prof. Hermann Oberth (right of the rocket) ready to test one of his first models with the student Wernher von Braun (second from right)

Another historical event happened on October 4, 1957, when the USSR successfully launched the *Sputnik 1*, the first artificial satellite, no bigger than a basketball, weighing 80 kg. Although this launch was an individual event, it provided the impetus the USA−URSS space competition (Box 6.2).

Box 6.2 Stages in journey to Space

Year	Platform	Novelty	Country
1920s	Rockets of Hermann Oberth	Pioneer in the Space	Germany
1944	V2: H. Oberth & W. von Braun	Ballistic missile beyond the Earth atmosphere	Germany
1957	Sputnik-1	Satellite in the Space	URSS
1961	Vostok-1	Gagarin in the Space	URSS
1969	LEM on Apollo 11	Armstrong and Aldrin on the Moon	USA
1972	ERTS-1 (Landsat-1)	Optical satellite for Earth Observation	USA
1978	Seasat	Radar satellite for Earth Observation	USA
1981	Space Shuttle	Reuse of the Space Shuttle	USA
1985	SPOT-1	Resolution 10 m	France
1991	ERS-1	Active microwave SAR	ESA
1998	International Space Station ISS	Permanent for experiments	International

On April 12, 1961 the first space shuttle, with the astronaut Yuri Gagarin on board, was launched from the space station Baikonur in the USSR. This was the beginning of the race to land men on the moon, reached by the LEM of Apollo 11 on July 20, 1969 with the astronauts Neil Armstrong and Edwin Aldrin, together with Michael Collins.

The age of modern remote sensing starts with the launch of the first artificial satellite *Earth Resources Technology Satellite* (ERTS) or Landsat-1, July 23, 1972.

Missions of other satellites followed: the *Space Shuttle*, since 1981, built to return to the atmosphere and reused, the French satellites SPOT, since February 22, 1986, with a better geometric resolution and stereoscopic nominal capacity, ESA ERS satellites with radar band instruments since 1991, up to the recent commercial satellites and the complex realization of the project of the *International Space Station* (ISS) started in 1998.

Several countries have rockets for launching satellites into orbit (Box 6.3). European members of the *European Space Agency* (ESA) developed the rocket *Ariane*, which performed the first launch in 1979 from Kourou's space centre, in French Guyana. After the first type, *Ariane 2* and *3* were produced in 1984, then *Ariane 4*, more powerful, in 1988, and *Ariane* 5, operative since 1999. The difference among these types is the capacity of the rocket of sent into orbit: *Ariane 2* and *3* allow payloads of 2 tons, *Ariane 4* more than 4 tons, *Ariane 5* nearly 7 tons. Hundreds of satellites have been launched into orbit by European rockets.

Box 6.3 Main carriers to launch in orbit satellites

Carrier	Country
Long March	China
Ariane 1–5 and Vega	Europe
PSLV and GLSV	India
H1 and H2	Japan
Proton and Soyuz	Russia
Atlas, Delta, Titan, Sea Launch, Pegasus	USA

Ariane 5 is the newest of the series and is different from previous versions. It is 54 m high and is made of three steps: the first one is 30 m long and 5.4 m in diameter, and its hydrogen and liquid oxygen engine, *Vulcain*, provides the rocket with the most propulsion. Two solid fuel propellants, laterally to the first piece, help in the first step; after burning for a couple of minutes, they are unhooked and fall down into the water. About 10 min after the launch, the engine of the first step is switched off and the second engine starts working: this one, burning mono-metil-hydrazine and nitrogen tetraoxide, works for 810 s and finally brings the payload into orbit. It is used to put into orbit two satellites at the same time and planned to send the automatic refill capsule towards the International Space Station (Box 6.4).

Box 6.4 Aeronautical and space definitions, characteristics and specifications

Term	Description	Specifics
Rocket	Endo-reactor operative out of the atmosphere	
Carrier	Rocket in different stages; the last one loads the satellite or hooks up the shuttle to be carried in orbit	
Satellite	Object in orbit around the Earth	Natural (Moon) artificial
Orbit	Path described by the rotation of a satellite around the Earth	Elliptic and ruled by the Kepler and Newton laws
Orbital parameter	Defining satellite position and velocity	Semiaxis Eccentricity Inclination Longitude ascending node Angle ascending node-perigee angle satellite-perigee
Orbit definition		Geostationary, Sun-synchronous or polar
Orbit inclination (i)	$i = 0°$ $i = 90°$ $0 < i < 90°$ $90° < i < 180°$	Equatorial polar direct retrogrades
Track	Projection of the satellite orbital path on the Earth's surface	

The European Space Agency is developing the Vega Programme, a new generation of rockets, for the launch of small and medium satellites, mainly commercial. Vega is a small rocket, about 27 m in height, 128 tons at launch (compared to *Ariane 5*'s 710 tons) planned to respond to the basic satellite launch requirement, intended for a low polar orbit, 90° inclined with respect to the equator, at 700 km altitude.

A curiosity is about the fact that when a satellite runs out of fuel, and then becomes unuseable, it is orbited out and lost in space. In case it is obsolete but potentially still operative, it is forced into a more internal orbit or to a different longitude, in order not to interfere with the other satellites, and reused in case of emergency or particular needs.

6.3.2 Satellite Platforms

Satellites for the Earth Observation (EO) are used mainly for:

- civil purposes;
- military strategies.

Recent evolutions of satellites for *military* use is little known, while their potential capacities of analysis and territorial survey, in terms of geometric, spectral, radiometric and temporal resolutions, are nowadays available for a wide range of applications.

In *civil* uses, they have two applications:

- *meteorological*: for weather forecasts (Meteosat and NOAA) and the study of the atmosphere (Terra, ENVISAT, ADEOS, METOP, etc.);
- *Earth resources*: automatic satellites, most cases (Terra, ENVISAT, Landsat, SPOT, ERS, IRS, Resurs, etc.), or with crew on board (*Space Shuttle*). The acquisition conditions are identical: usually nadir, from platforms at near-polar and sun-synchronous orbit.

Satellite platforms planned for Earth Observation are in continuous development. After the relevant political–economical changes at global scale, also Russia, at the end of the 1980s, liberalized the market of satellite images and photographs taken from space (the latest acquired in the long space missions with crew, previously rigorously classified by the military). Besides Russia, USA, France, Japan and European Community, as well as Canada, India, China, Brazil, Argentina, Israel, Germany and Italy are actively taking part in the development of satellite platforms with their own sensors and technology and many other countries cooperate in the development of international programmes. The liberalization of restrictions, especially geometrical, has led to the development of high-resolution commercial satellites.

Satellite remote sensing offers operators relevant opportunities for both the objects' identification and surface dynamic process analysis.

Images acquired by different instruments-sensors systems are increasingly required (Box 6.5), each one more suitable to recognize and develop a particular aspect of the problem and, all together, enabling a synergy otherwise not possible.

First-generation satellites and commercial satellites carried on board one or two instruments; complex programmes for experimental scientific missions carry several instruments. The largest contribution to terrestrial surface monitoring has been sequentially offered since the early 1970s by US satellite systems such as Landsat and NOAA, followed by the French SPOT, the European ERS-1 and 2, the Canadian Radarsat, the Indian IRS 1C and 1D, equipped with advanced optical–mechanical or electronic acquisition systems. They faced specific issues and independently, sometimes with the same purposes but with scarce or no interaction and cooperation.

Box 6.5 Some important instruments on board satellites for the study of the atmosphere and of the terrestrial surface

Sensor	Resolution	Sensor objective
Visible and near infrared radiometer passive 0.4–1.1 μm	0.4–1100 m	Radiance measurement of land–water elements
Medium IR radiometer passive 1.4–2.5 μm	1–1100 m	Geological studies
Thermal IR radiometer passive 3–30 μm	15–1000 m	Superficial temperature measurement
Microwaves radiometer passive	4–75 km	Radiance measurement of land–water elements in the microwaves
Microwaves radar active	1–100 m	Measurement of backscattering signals of land–water elements
Radar altimeter	< 1 m	Measurement of objects height on the Earth's surface Measurement of glaciers thickness
Laser altimeter	10–30 cm	Height measurement of objects on the Earth's surface
Scatterometer	6–50 km	Velocity and wind measurement radiative balance Passive sounding of the atmosphere Chemistry of the troposphere
Tropospheric Wind Lidar	0.5–40 km	Laser technology at solid state

Instrument characteristics are defined during the planning phase, in relation to the data to be collected with, generally, well-defined objectives.

This book deals with sensors acquiring images of the land–water surface, i.e. satellites for Earth Observation (EO) studying terrestrial resources. Sensors for studying the solar and planetary systems and those studying the Sun and atmospheric behaviour and phenomena are omitted here.

EO satellites have a speed of over 30,000 km/h and do not need engine propulsion over the rarefied or non-existent atmosphere (400 km). The presence of little quantities of ionized oxygen and nitrogen generates a slight resistance to the satellite's motion, continuously requiring both the orbit and the speed to be known and the periodical corrections by on board small propellants. This correction is necessary to circumvent the Earth's gravitational field attraction.

Acquisitions can be monoscopic, or stereoscopic when at least two images with different angles of view are acquired; the acquisitions can be along the flight line, or with a lateral view pointing the instrument on the same scene during an orbit which follows the reference one (Fig. 6.20).

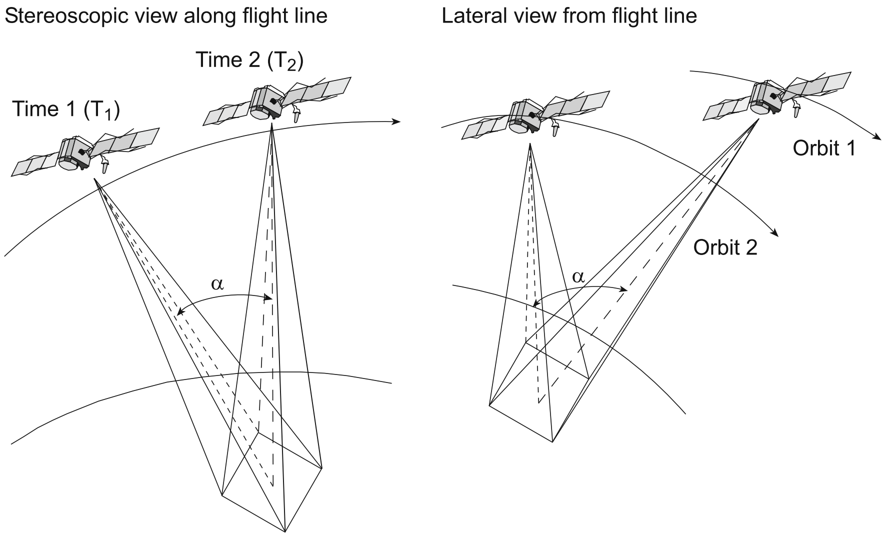


Fig. 6.20 Stereoscopic acquisition along the flight line; (**a**) on the same orbit, (**b**) lateral vision from a different orbit

6.3.2.1 Spectral Intervals

Before the description of space programmes and satellites, the definition of *spectral intervals*, or *bands*, or *channels* characterizing the most commonly used satellites sensors operating in the optical reflected and emitted wavelengths (Figs. 6.21 and 6.22) is introduced, while frequencies in the range of the microwaves are treated in Chapter 4:

- *0.50–0.75* μ*m (panchromatic-Pan-P)*: with low spectral resolution, it gives an important contribution to the study of images, having geometric resolution at least four times higher than the equivalent multispectral bands;
- *0.45–0.52* μ*m (blue−green)*: spectral range used for studying water transparency, thanks to its limited but significant penetration in water bodies. About vegetation study, it detects changes in the chlorophyll/carotenoids pigments ratio;

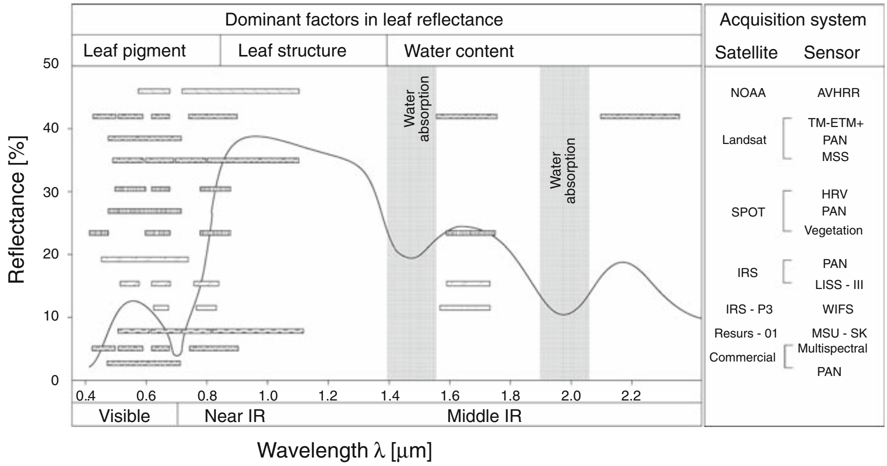


Fig. 6.21 Visible, near and middle infrared spectral ranges of the main satellite sensors. On the backdrop, the vegetation spectral response

- *0.52–0.60* μ*m (green)*: this interval is used to measure the peak of reflection of vegetation in the wavelength corresponding to the green for determining its vigour condition. For water bodies, the ratio to the blue−green band gives information about plankton and suspended organic matter;
- *0.63–0.69* μ*m (red)*: basic band to discriminate between vegetation classes, due to the different chlorophyll absorption values among the species and to the sensitivity to tannins and anthocyanins content. The spectral behaviour differences between vegetated areas and bare soils are highlighted. The atmosphere effect is low, so the surface shapes appear sharper;
- *0.76–0.90* μ*m (near or photographic infrared)*: used for studies about biomass and water content, in this region the maximum values of reflection for vegetation are recorded. Particularly suitable to distinguish water bodies and hydrographical network;
- *1.55–1.75* μ*m (medium infrared)*: leaves' biomass reflection in these intervals is related to its water content; this band is then useful to detect vegetation water stress, linked to the leaf turgidity. It is also useful to discriminate soils with different water content, drainage conditions and organic matter content. Used for distinguishing between clouds, which absorb and appear dark, and snow cover, which reflects and looks lighter;
- *2.08–2.35* μ*m (medium infrared)*: used for its potentiality in distinguishing lithotypes, particularly in areas where rocks are subjected to hydrothermal alterations;
- *3.55–3.93* μ*m (medium/thermal infrared)*: for studies about water vapour in the thermal range, this spectral range is still affected by the radiation reflection;
- *10.4–12.5* μ*m (far or thermal infrared)*: used to assess the evapotranspiration of vegetated areas. Suitable for studies about stress related to the increase of temperature in the biomass. The radiation emitted by the surface is linked with the emissivity and its absolute temperature; it can be used for studying the thermal balance and the thermal inertia.

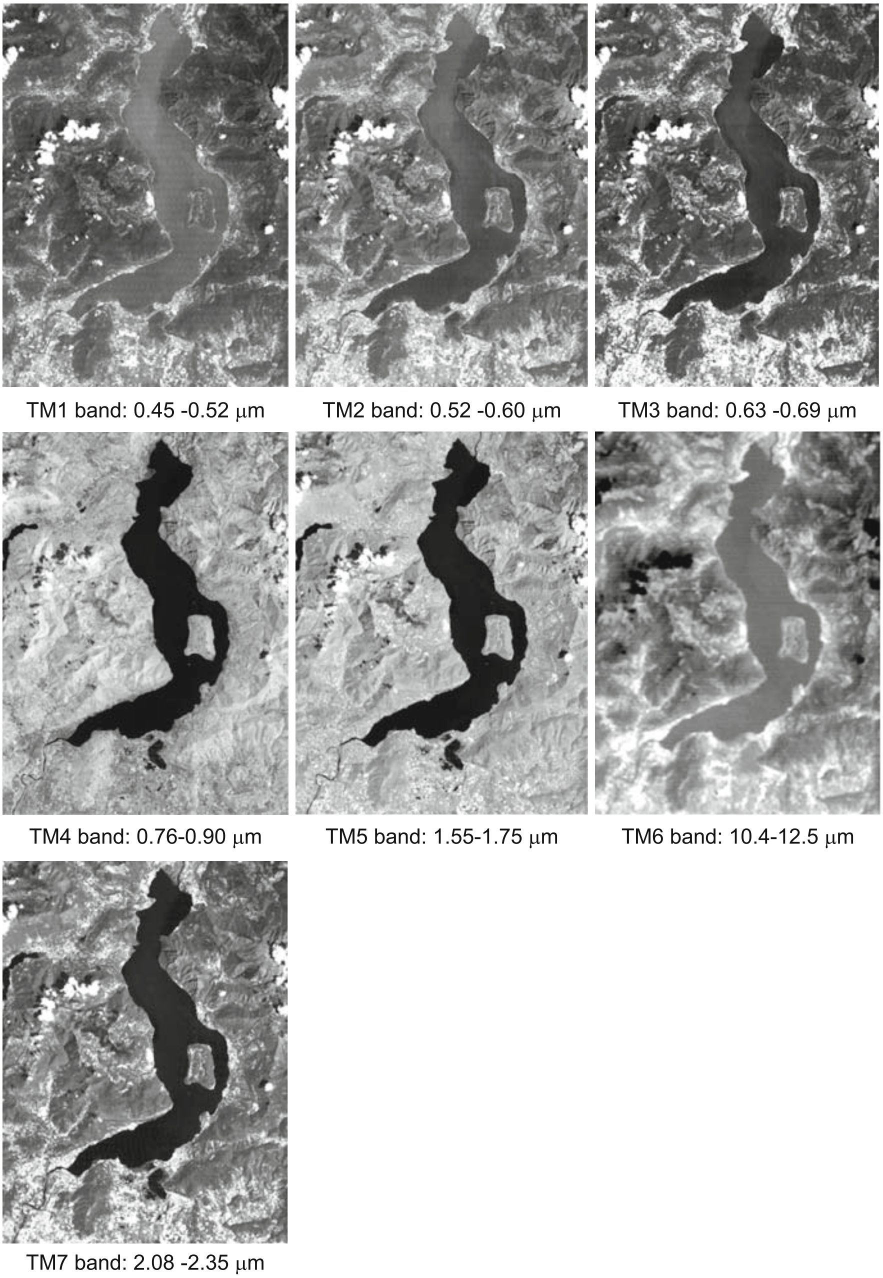


Fig. 6.22 The seven Thematic Mapper (TM) spectral bands of the satellite Landsat-5, scenes 194/28, Iseo Lake, Northern Italy

6.4 Earth Observation Space Programmes

The development of satellite technology is fundamental for studying, observing, monitoring and assessing terrestrial resources. The atmosphere–earth–water complex system must be studied in its whole, i.e. in *global* terms, in order to better understand phenomena otherwise not observable.

Since mid-1980s, the EO system has increased their importance, currently organized through the *Integrated Global Observing Strategy* (IGOS), which uses the main remote and ground sensing systems for the global observation of the atmosphere, the sea and the land aiming both to understand and monitor natural processes, and to assess human activities impact.

Organisms in IGOS also include

- *Committee on Earth Observing Satellites* (CEOS);
- integrated research programmes like the *World Climate Research Programme* (WCRP) and the *International Geosphere–Biosphere Programme* (IGBP);
- many international agencies like FAO, UNESCO and *World Meteorological Organisation* (WMO);
- *Global Climate Observing System* (GCOS), *Global Ocean Observing System* (GOOS) and *Global Terrestrial Observing System* (GTOS), which organize global scale observations in operative terms.

The terrestrial components (vegetation, water, soil) of the Earth system are not intended as passive elements, but are recognized as active elements of the Earth−atmosphere system. In this new perspective, the accurate representation of the dynamics of surface phenomena (land cover change) and exchange processes (mass and energy fluxes) assume a more important role for climate modelling and global changes.

This is the reason why the leading Space Agencies (Box 6.6) worldwide develop programmes based on complex, dynamic and interdisciplinary scenarios that may allow both continuity in acquisition and measurement of the most important geophysical and biophysical parameters and of phenomena globally studied and evaluated.

The common effort is to develop systems from a qualitative approach to a perspective of quantitative observations which can be assimilated into models. Practical collaborations were adopted following the meeting of the G8 Nations at Evian, France, in 2003, creating the *Group of Earth Observation* (GEO), tasked with conceiving a worldwide Earth observation programme: the *Global Earth Observation System of Systems* (GEOSS). This programme was adopted at the 3rd Earth Observation summit in Brussels, in February 2005 (see Chapter 9).

Wide applicative opportunities for a large range of both private and public users can be suggested. The users need to have the availability of data about the terrestrial surface condition also in relation to international conventions like the Kyoto Protocol (1997). Following ratification by Russia, the Kyoto Protocol entered into force on February 16, 2005.

Box 6.6 Leading Space Agency in the world

Acronym	Space Agency	Country
CONAE	Comision National de Actividades Espaciales	Argentina
AEB	Agência Espacial Brasileira	Brazil
CSA	Canadian Space Agency	Canada
ESA	European Space Agency	15 European Countries
CNSA	China National Space Administration	China
CNES	Centre National d'Etude Spatiales	France
DLR	Deutschen Zentrum Fur Luft-und Raumfahrt	Germany
NRSA	National Remote Sensing Agency of India	India
ISA	Israeli Space Agency	Israel
ASI	Agenzia Spaziale Italiana	Italy
JAXA (NASDA)	Japan Aerospace Exploration Agency	Japan
KARI	Korea Aerospace Research Institute	South Korea
FSA (RASA-RKA)	Federal Space Agency (Roskosmos)	Russia
BNSC	British National Space Centre	United Kingdom
NASA	National American Space Administration	USA
NOAA	National Oceanic and Atmospheric Administration	USA

6.4.1 EUMETSAT: Geostationary Meteorological Satellites Network

EUMETSAT (*European Organisation for the Exploitation of Meteorological Satellites)* is an inter-government organization, whose programmes are promoted and financially sustained by 17 European nations: Austria, Belgium, Denmark, Finland, France, Germany, Great Britain, Greece, Ireland, Italy, Norway, Holland, Portugal, Spain, Sweden, Switzerland and Turkey. These countries are the main users of this system, which carries out the following activities:

- launching and monitoring the correct functioning of in orbit satellites;
- meteorological data dissemination to the end users;
- climate continuous monitoring, for changes forecast and adequate prevention.

EUMETSAT has been established in 1986, after the launch of the first Meteosat satellite in 1977 by the European Space Agency (ESA), which was successfully used for 2 years. In November 1979 Meteosat-1 radiometer, which had acquired excellent quality images every 30 min, for a technical breakdown stopped the activity. The work done by Meteosat-1 was considered indispensable and the total success of the mission led to an intergovernmental meeting after which EUMETSAT was founded, undertaking the Meteosat satellites programme from ESA.

In January 1987 Meteosat-3 replaced Meteosat-2 launched in 1981; the generation ended with Meteosat-7. After those, in sequence, the Meteosat Second Generation (MSG) satellites followed, in geostationary orbit, 0° latitude (equator) and 0° longitude (Greenwich meridian) (see Plate 6.2).

Meteosat satellites have 2.5 km resolution in the visible and 5 km resolution in the infrared. MSG satellites are able to provide better details and 20 times more data.

They belong to the worldwide network of meteorological satellites positioned at 36,000 km over the equator at a distance of 70° longitude one from the other in order to optimize the acquisition and availability of frequency data of the world. The other satellites constituting the network are the US GOES series, the Japanese GMS and the Russian series.

EUMETSAT Polar System (EPS) took place in 1994 as complementary programme of the geostationary system in order to acquire higher resolution images. Polar orbits, below 1000 km of altitude, provide more data and details about atmosphere temperature and humidity profiles, although with a lower frequency of acquisition. EUMETSAT instruments are on board the METOP satellites, developed in cooperation with ESA.

The Meteosat system is intended to support the National Meteorological Services. There are about 2000 systems, spread across 100 nations, used for real-time reception of images from the satellites controlled by EUMETSAT.

6.4.2 NOAA Meteorological Programme

NOAA satellites were born in the 1970s as meteorological, low (about 850 km) sun-synchronous orbit satellites. They had success in global change studies. They are efficiently used measuring the sea surface temperature, studying the snow cover, monitoring vegetation conditions and analysing thermal inertia.

The first polar orbit meteorological satellite TIROS (*Television Infrared Observation Satellite*) was launched on April 1, 1960 by NASA's *Goddard Space Flight Center* from Cape Canaveral, USA. It offered the chance to monitor the cloud cover at global scale. From April 1960 to 1965, 10 TIROS satellites were launched. Up to 1970 another two series of ESSA satellites (*Environmental Science Service Administration*) and ITOS (*Improved Tiros Operational Satellite*), precursor of NOAA series, followed.

In October 1970, after the establishment of NOAA (*National Oceanic and Atmospheric Administration*), the first NOAA-1 satellite was launched, operative until August 19, 1971.

The sensor AVHRR (*Advanced Very High-Resolution Radiometer*) was launched for the first time in 1979: it has five spectral bands respectively in the red, near infrared, medium infrared and two in the thermal infrared (Table 6.2). AVHRR represents the core of the sensors set on board NOAA satellites.

Thanks at AVHRR-1 resolution, 1.1 km (it can reach 0.5 km in AVHRR-2 and 3), one scene covers a very large area (2600–2940 km^2) with high temporal resolution: NOAA satellites pass above the same area twice a day every 12 hours. Considering that at least two NOAA satellites, alternated both in the orbits and in the time of passing, are always in orbit, four scenes of the same area can be obtained per day. The odd series (NOAA-11, 13, 15, 17, 19 ...) passes at 2 a.m. and 2 p.m. while the even one (NOAA-10, 12, 14, 16, 18 ...) at 7 a.m. and 7 p.m.

Table 6.2 Instrumental characteristics of sensors on board NOAA satellite series

Acronym	Sensor	Geometric resolution (m)	Swath width (km)	Spectral bands (μm)
AVHRR	Advanced very high-resolution radiometer	1100	2940	0.58–0.68 0.72–1.10 3.44–3.93 10.3–11.3 11.3–12.5
AVHRR-2	Advanced very high-resolution radiometer 2	500 visible 1000 infrared	2600	as AVHRR
AVHRR-3	Advanced very high-resolution radiometer 3	500 visible 1000 infrared	2940	0.58–0.68 0.72–1.00 1.58–1.64 day 3.55–3.93 night 10.3–11.3 11.5–12.5

6.4.3 NASA (USA) Space Programme

The NASA has constituted the *Earth Science Enterprise* (ESE) which, besides management, planning and coordination aspects, research and applications, dissemination and training, includes the *Earth Observing System* (EOS) space sector. It is committed to technological development of satellites, instruments and sensors and caring about data calibration and validation. EOS includes Landsat, EOS-AM Earth and Water programmes and also has links with activities of other space agencies (Box 6.7).

Box 6.7 Some satellites, instruments and objectives of the Earth Observing System (EOS) programme

Satellite	Sensor	Mission objective
Landsat 7	ETM+	Monitoring of the Earth's surface with measuring of radiance/reflectance in visible, infrared and thermal infrared at medium-high spatial resolution
EOS-AM Terra	MODIS MISR ASTER CERES MOPITT	Characterization of the terrestrial ecosystems, land use/land cover, soil moisture, terrestrial energy, oceans primary productivity, characteristics of the clouds systems, aerosol and radiative balance
EOS-AM Aqua	MODIS AMSR-E CERES AMSU	Besides MODIS, of interest AMSR-E, microwaves radiometer to study the water cycle and the radiative energy flux

EOS-AM Aura	HIRDLS MLS OMI TES	Chemical composition of the atmosphere and ozone thickness monitoring and gas pollution of the planet in synergy with Calipso and Parasol.
CALIPSO	CALIOP WFC IIR	Aerosols and clouds measurement 24 h a day. CALIPSO will fly in constellation with several other satellites: Aqua, Aura, CloudSat and the French PARASOL. Wide Field Camera selected to match band 1 of the MODIS instrument on Aqua Detection of cirrus cloud emissivity and particle size.
PARASOL	POLDER	Radiative and microphysical properties of clouds and aerosols
ENVISAT series	MERIS AATSR ASAR	Environmental studies, chemistry of the atmosphere, marine biology and prosecution of the ERS missions objectives: SAR in band-C, altimeter, scatterometer, ocean surface temperature
ESSP/VCL	MBLA GPS	Measurement of the height of the vegetation cover using the Laser Altimeter (VCL) and topographic studies of the Earth's surface with resolution <1 m (GPS)
ADEOS-II	GLI POLDER AMSR SeaWinds	Radiance/reflectance from visible to thermal, images in radar band, scatterometry, measurement of atmospheric parameters, study of snow and ice, marine biology, oceans stream and colour
METOP series	AVHRR-3 ASCAT . . .	Operative for the climate monitoring and the meteorology with the future objective of operativeness also in climatology, combined with the NOAA satellite series
Hyperion	HSI	Hyperspectral sensor with 220 spectral bands and 30 m geometric resolution for environmental studies.

In order to get the most data about the hydrological cycle, NASA decided to launch a series of space platforms and small satellites having different sensors on board, to cover the greatest part of the electromagnetic spectrum, and thus able to collect as much data as possible about geophysical parameters of solid Earth, sea and atmosphere. After EOS-AM Earth, Water and Aura, launched respectively in 1999, 2002 and July 15, 2004, followed PARASOL, launched December 18, 2004, carrying an instrument called POLDER that studies the radiative and microphysical properties of clouds and aerosols, and CALIPSO *(Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations*) April 28, 2006.

6.4.3.1 Landsat Series

Landsat satellites operate on circular, near-polar, sun-synchronous orbits, at altitude of about 920 km (1st generation) and 705 km (2nd generation).

Within NASA ERTS (*Earth Resources Technology Satellite*) programme, seven satellites were launched (Box 6.8).

Landsat 1, 2 and 3 first generation complete an orbit in 103 min, thus completing 14 orbits a day and scanning the Earth's surface in 18 days. Landsat 4, 5 and 7 have orbits covering every 16 days the same area crossing the equator at 10:30 a.m., solar time (see Plate 6.3).

Box 6.8 Landsat satellites series

Satellite	Launch date	Operativeness
Landsat 1 (ERTS-1)	July 23, 1972	January 6, 1978
Landsat 2	January 22, 1975	July 27, 1983
Landsat 3	March 5, 1978	September 7, 1983
Landsat 4	July 16, 1982	November 1996
Landsat 5	March 1, 1984	Operative
Landsat 6	October 5, 1993	Lost in the space in preliminary orbit correction phase
Landsat 7, copy of Landsat 6	April 15, 1999	Operative

Landsat 1, 2 and 3 carried two multispectral systems on board:

- *Return Beam Vidicon* (RBV): a special type of vidicon tube, with three spectral bands green, red and near infrared, using the electrons beam to create images reproduced in a video signal with 40 m × 40 m ground resolution;
- *Multispectral Scanner* (MSS): operative also on Landsat 4 and 5, characterized by four spectral bands with geometric resolution of 79 m × 79 m. The same bands were later used in MOS-1 Japanese satellite and Resurs-01 Russian satellite with different geometric resolutions (50 m and 160 m respectively).

Landsat 4 and 5 satellites carry the *Thematic Mapper* (TM), Landsat 7 the *Enhanced Thematic Mapper plus* (ETM+), characterized by seven spectral bands.

The main characteristics of TM are

- Rotating mirror around an axis parallel to the flight direction, producing a transversal scanning of the scene. The oscillation frequency is 7 Hz, i.e. a period of 142 ms. The acquisition is carried out during the mirror semi-oscillation from west to east as well as during the return phase; the ground area explored during each of the two oscillation phases is 185 km (100 miles) wide, approximately north–south oriented;
- 16 scan lines are swept simultaneously for the spectral bands 1, 2, 3, 4, 5 and 7, with IFOV of 30 m × 30 m, while band 6 is scanned by four detectors with 120 m × 120 m ground resolution.

The detectors of bands 1–4 are constituted by silicon photodiodes, bands 5 and 7 detectors by indium antimonite photodiodes and band 6 by cadmium and mercury telluride.

Collected data are codified into binary numbers series by an analogical−digital converter and transmitted to a ground station. Data can be recorded on a tape on board the satellite and transmitted to the receiving station when the satellite passes in the reception range, or sent to TDRSS (*Tracking Data Relay Satellite Systems*) geostationary satellites which forward the data to the ground station.

The dynamic range is of 8 bit (256 Digital Numbers).

6.4.3.2 Landsat 7

After the launch of Landsat-6 failed on 5 October 1993, on 15 April 1999 Landsat 7 was successfully launched, on a circular, near-polar, sun-synchronous orbit, at 705 km altitude. This generation of Landsat improved technical devices maintaining the same basic characteristics (swath width of 185 km and 16 day temporal resolution), which ensure the continuity of Landsat missions since 1982. It is equipped by 1 panchromatic (Pan), 6 multispectral (MS) reflective bands and 2 thermal infrared (TIR) bands.

Compared to Landsat 4 and 5, Landsat 7 has an updated version of the sensor TM, the *Enhanced Thematic Mapper Plus* (ETM+), which introduces the following improvements:

- 15 m geometric resolution panchromatic band (Fig. 6.23), extended from the visible to the near infrared in the spectral range 0.52–0.90 μm, acquired simultaneously with the other multispectral bands and thus co-recorded with them. The pixel of 15 m permits the overlapping, in 4 to 1 ratio, with the 30 m multispectral pixel. This solution remains within the limit of data flux of 85 Mbps (Megabytes per second) that can be downloaded from the Landsat 7;
- 60 m resolution thermal band, in the interval 10.40–12.50 μm, available in high and low restitution versions;
- better *signal-to-noise ratio* (SNR) with in flight solar calibration that allows an improvement of the radiometric accuracy up to 5%;
- larger capacities for data storage in two recorders (*Landsat Recorder System*, LRS), each one having 15 min acquisition time equal to 29 scenes, which can be downloaded when the satellite passes near a receiving ground stations;
- automatic system for cloud covers assessment that prevents from the acquisition and storage of very cloudy images.

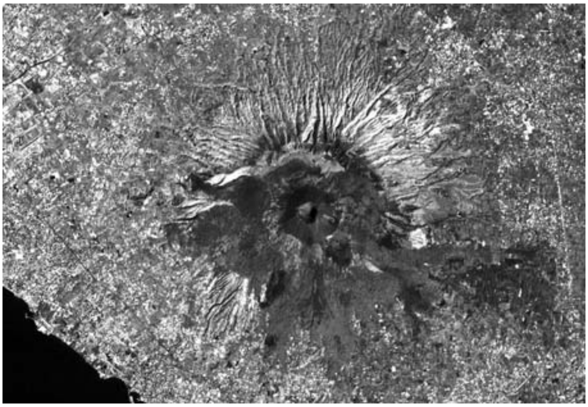


Fig. 6.23 Landsat-7 panchromatic band, 15 m geometric resolution, Vesuvio Volcano, Southern Italy

The introduction of the panchromatic band is the most important innovation; covering the visible range (0.5–0.7 μm) as well as part of the near infrared (VIR: 0.7–0.9 μm), a better signal-to-noise ratio is obtained, improving the image readability.

Fusion techniques (pan-sharpening) permits the integration of the panchromatic with the multispectral bands obtaining true and/or false-colour images with 15 m geometric resolution (see Chapter 8).

The simultaneous acquisition of panchromatic and multispectral images with homogeneous characteristics from the same satellite platform allows co-registration, time saving and processing costs reduction.

The introduction of a thermal band with better geometric and radiometric characteristics allows the enlarging of the range of applications about natural or anthropic activities, such as heat exchange, volcanic phenomena, pollution, thermal emissions.

The improved geometric resolution (15 m) and the confirmed temporal acquisition (16 days) can give a contribution to global change assessment, both providing elements for verifying land cover changes and defining local processes that can induce these modifications. Applications can include studies about deforestation, ecosystems fragmentation, crops production, glaciers dynamics, coastal monitoring, volcanoes monitoring, etc.

6.4.3.3 EOS-AM Terra

This satellite was successfully launched on December 18, 1999 into a polar, sun-synchronous orbit at 705 km altitude. It carries five different instruments collecting information about physical and radiation properties of both the atmosphere and the terraqueous system (Box 6.9). Fifteen satellites of the EOS programme are scheduled to be launched in the next years. The purpose of the programme is to discover more about climatic and environmental changes.

Box 6.9 Instruments, acronyms and field of operability of the devices on board the satellite EOS-AM Terra

Acronym	Sensor	Sensor objective
MODIS	*Moderate resolution imaging spectroradiometer*	Study of the system atmosphere–water–land
MISR	*Multi-angle imaging spectroradiometer*	Multi-angle detection of the same scene
ASTER	*Advance spaceborne thermal emission and reflection radiometer*	Japanese multispectral device with spatial resolution from 15 m to 90 m; for studying climatic processes, clouds and land.
CERES	*Clouds and the Earth's radiant energy system*	Accurate measurement of clouds and radiative flux, fundamental elements for atmosphere and oceans modelling and for meteorological forecasting
MOPITT	*Measurements of pollution in the troposphere*	Infrared radiance measurement reflected and emitted in an atmospheric column, quantifying the content of carbon oxides (CO) and methane (CH_4).

EOS-AM Terra satellite, like all this type of satellites, has a descending orbit that crosses the equator at 10:30 a.m., a time of the day with minimum cloud cover.

The main instrument on board is *MODIS* (*Moderate Resolution Imaging Spectroradiometer*). It has characteristics similar at Envisat-MERIS, EOS-AM Aqua, and others. MODIS is planned for measuring parameters related to biological and physical processes at global scale every 1 or 2 days. The high temporal resolution permits the acquisition of more cloud-free images.

MODIS is a multidisciplinary instrument planned to know more about the complex terrestrial system providing data about:

- *atmosphere*: aerosol and clouds properties, water vapour and temperature profiles;
- *seas and oceans*: surface temperature, fluorescence and chlorophyll concentration;
- *terrestrial surface*: identification of land cover classes and study of temporal changes, surface temperature, snow cover assessment and reflectance, vegetation properties such as net primary production, *Leaf Area Index* (LAI), vegetation indices, soil and radiation direction effects, study of the photosynthetically active radiation.

The acquisition system adopts solutions typical of transversal oscillating scanning conventional mirror and optical spectroradiometers combined with electronic sensors (linear array). The system acquires images with 36 spectral bands in the interval from 0.4 to 14.5 μm, selected to have diagnostic meaning in the study of terrestrial resources. Geometric resolution can be changed in 250 m (2 bands at a time), 500 m (5 bands), and 1000 m (29 bands) at nadir. The signal-to-noise ratio is favourable to the signal, and the absolute radiance accuracy is lower than ± 5% from 0.4 to 3.0 μm and <1% in the thermal infrared interval 3.7–14.5 μm. MODIS regularly provides day time reflection data and day/night emission data with 2 days temporal resolution. In the second case, the following data are collected:

- sea and ocean surface temperature with 0.3–0.5 K accuracy;
- terrestrial surface temperature with 1 K accuracy.

MISR (*Multi-angle Imaging SpectroRadiometer*): can acquire the same scene from different angles in a few minutes. It is equipped with nine CCD (*Charge-Coupled Device*) digital cameras observing from nine angles of view: vertical, and four symmetric acquisitions at 26.1°, 45.6°, 60.0° and 70.5° along the line of acquisition. Each camera can acquire in four spectral bands centred in 0.446, 0.558, 0.672 and 0.886 μm wavelengths. Each of the 36 detectors (nine digital cameras with four spectral bands each) can be individually programmed with the possibility to obtain 240, 480, 960 m or 1.92 km geometric resolution. The swath width of 360 km enables the multi-angle global acquisition in a period between 2 and 9 days. The adopted devices maintain the absolute radiometric uncertainty in the interval ±3% for surfaces lightened and ± 6% for shadowed surfaces.

MISR images can be acquired in global or local way.

Table 6.3 Characteristics of the instrument ASTER on board the EOS-AM Terra satellite

Spectral region	Acronym	Number of bands	Spectral range (μm)	Geometric resolution (m)	Temporal resolution (days)	Stereo
Visible and near infrared	VNIR	2	0.52–0.6	15	4	No
		1	0.63–0.69	15	4	Yes
			0.76–0.86			
Medium infrared	SWIR	6	1.6–2.43	30	16	No
Thermal infrared	TIR	5	8.12–11.65	90	16	No

The global acquisition allows the continuous observation of the terrestrial surface with most of the available bands operating at medium or low resolution (960 m or 1.92 km) and with some selected bands operating at the best available resolution of 275 m. This acquisition acts for cloud cover mapping and classification, image navigation and small scale stereo-photogrammetric applications.

The local acquisition acquires data of 240 and 480 m resolution in 36 spectral bands over a sub-window of 300 km swath width that can be selected within the field of view (FOV).

This tool provides different measures related to the terrestrial surface such as bi-directional (BRDF) and hemispheric reflectance, Leaf Area Index (LAI) and identification of land cover classes. Used in combination with MODIS, it can improve land use statistical and dynamic classification, helping in investigating photosynthesis and biomass transpiration aspects, at scales compatible with the geometric resolutions.

ASTER (*Advanced Spaceborne Thermal Emission and Reflection Radiometer*): multispectral scanning device operating at variable spatial resolutions, from 15 to 90 m, in 14 different spectral bands in visible, near and thermal infrared. The swath width is 60 km. ASTER characteristics are reported in Table 6.3.

In the visible and near infrared, the temporal acquisition varies from 4 to 16 days due to the instrument's capability of off-nadir pointing at ±24°. The stereoscopic capacity, due to the presence of an additional telescope acquiring in the same spectral range of the band 3 (NIR 0.7–0.9 μm), generates the Digital Elevation Model (DEM) with 7 m horizontal resolution, if supported by the acquisition of Ground Control Points (GCPs), and 1 m vertical resolution. The possible applications are in land use/land cover multi-temporal classification, geological studies, volcano monitoring, surface temperatures, surface emissivity and reflectivity determination, and evapotranspiration assessment.

The other instruments on EOS-AM Terra, CERES and MOPITT, are focused on cloud monitoring, radiative flux and atmospheric components.

6.4.3.4 EOS-AM Aqua

Aqua is a satellite so named for its specific mission concerning water cycle and climatic global changes.

Aqua has been developed as a joint project by the United States, Japan and Brazil.

This satellite, launched on May 2, 2002, is equipped with the MODIS sensor, the same as on board EOS-AM Terra and with a microwave multi-band radiometer of new generation: AMSR-E (*Advanced Microwave Scanning Radiometer*) identical to the one on board the Japanese platform ADEOS-II. This radiometer, operating in the domain of passive microwaves, has advanced characteristics compared with the previous generation and in particular it has a better ground resolution, although it is still of the order of tens of kilometres, especially at lower frequencies. It is suitable for studying phenomena related to global scale water and energy cycles: ocean evapotranspiration, atmospheric water vapour, clouds, precipitation, soil moisture, terrestrial and sea glaciers, land cover, phytoplankton and organic compounds dissolved in the oceans, atmospheric temperatures, soil and water.

6.4.3.5 EOS-AM Aura

Aura (formerly EOS Chemistry-1) is part of the EOS programme, dedicated to monitoring the complex interactions that affect the Earth.

The Aura satellite, launched July 15, 2004 from the Western Range of Vandenberg Air Force Base, California, into 705 km orbit, focuses on measurements of atmospheric trace gases and their transformations. The objective of the mission is to study the chemistry and dynamics of the Earth's atmosphere from the ground through the mesosphere.

The Aura spacecraft has on board the following instruments:

- *High-Resolution Dynamics Limb Sounder* (HIRDLS)
- *Microwave Limb Sounder* (MLS)
- *Ozone Monitoring Instrument* (OMI)
- *Tropospheric Emission Spectrometer* (TES)

The main goals assigned to these instruments are

- monitoring levels of ozone in the stratosphere, which protects life from the harmful effects of ultraviolet radiation from the Sun, and the pollutants, that contribute to depleting it;
- space-based observations of the troposphere to discover global patterns and trends that can affect human activity;
- studying the levels and distribution of water vapour and ozone in the upper layers of the atmosphere to better understand how they regulate climate.

6.4.3.6 Earth Observing-1 (EO-1) Satellite

EO-1 satellite is part of the NASA programme; it was launched on November 21, 2000. After the initial technology mission was completed, NASA and USGS agreed to the prolongation of the EO-1 programme as an *Extended Mission* to collect and distribute Hyperion hyperspectral and *Advanced Land Imager* (ALI) multispectral products according to customer tasking requests.

Hyperion is based on the experience of hyperspectral missions of the LEWIS project and provides hyperspectral data in 220 bands in the spectral range from

0.357 to 2.576 μm with 10 nm bandwidth and 30 m geometric resolution. The standard scene width is 7.7 km. The acquisition has high radiometric accuracy that generates a very good signal-to-noise ratio (SNR).

This mission gives the start to medium resolution hyperspectral study from space with wide possibilities of complex environmental analysis.

ALI provides image data from 10 spectral bands. The CCD pushbroom instrument has a spatial resolution of 30 m, multispectral bands, and 10 m, panchromatic band. The standard scene width is 37 km. Scene length is 42 km, with an optional increase to 185 km, as for Hyperion.

6.4.4 ESA (European Union) Space Programme

The European Space Agency (ESA) has proposed the *Living Planet* programme, which affects European and global investments and choices about space studies and applications for the next 25 years.

The aims of these programmes are very ambitious and global; they concern generic data acquisition, like

- implementing knowledge about the Earth;
- preserving the Earth and its environment;
- efficiently managing life on the Earth.

These themes are approached in two ways:

- *Earth Explorers*, having scientific and research elements and demonstration objectives, realized by missions planned for better understanding processes characterizing our planet;
- *Earth Watch* element, designed to facilitate the delivery of Earth Observation data for the eventual use in operational services. Earth Watch includes the meteorological missions within EUMETSAT and missions focusing on the environment and civil security under GMES (*Global Monitoring for Environment and Security*) a joint initiative of European Commission and ESA.

The Earth Explorer missions encompass a new strategy for Earth observation where missions are designed to address critical and specific issues that have been raised by the scientific community, introducing a peer-reviewed selection process, whilst demonstrating breakthrough technology in observing techniques.

The family of Earth Explorer includes six approved missions and several candidate missions for Earth Observation:

- GOCE (*Gravity field and steady-state Ocean Circulation Explorer*) will provide the data set required to accurately determine global and regional models of the Earth's gravitational field and geoid. It will advance research in areas of ocean

circulation, physics of the Earth's interior, geodesy and surveying, and sea-level change.
- SMOS (*Soil Moisture and Ocean Salinity*) will provide global maps of soil moisture and ocean salinity to further our understanding of the Earth's water cycle and contribute to climate, weather and extreme-event forecasting.
- ADM-Aeolus (*Atmospheric Dynamics Mission*) will make novel advances in global wind-profile observation and will provide much-needed information to improve weather forecasting. This mission should pave the way for future operational meteorological satellites dedicated to measuring the Earth's wind fields.
- CryoSat-2 will determine variations in the thickness of the Earth's continental ice sheets and marine ice cover to advance our understanding of the relationship between ice and global warming. CryoSat-2 replaces CryoSat, which was lost at launch in 2005.
- Swarm: a constellation of three satellites to study the dynamics of the magnetic field to gain new insights into the Earth's system by improving our understanding of the Earth's core and climate.
- EarthCARE (*Earth Clouds Aerosols and Radiation Explorer*): a joint European–Japanese mission that aims to improve the representation and understanding of the Earth's radiative balance in climate and numerical weather forecast models.

Among the candidate missions for Earth observation, there is the *Land-Surface Processes and Interactions Mission* (LSPIM). This mission, named as the satellite, aims to measure terrestrial surface characteristics such as albedo, reflectance, *Bi-directional Reflectance Distribution Function* (BRDF), surface temperature, factors related to soil bio-geophysical and geochemical processes. The *Processes Research by an Imaging Space Mission* (PRISM) hyperspectral sensor, pushbroom-based imaging spectrometer, is going to be developed and placed on board the LSPIM: it is supposed to record the same scene in 142 different bands in the visible, near and medium infrared with 10–15 nm spectral resolution. The system is completed by a two spectral band radiometer acquiring in the thermal infrared. The geometric resolution is 50 m $\times$ 50 m with a swath width of 50 km at nadir. The instrument can also point along the swath or laterally for observations and studies of the BRDF with 3 days acquisition period.

In the context of the Living Planet Programme and along with the Galileo satellite navigation programme (see Chapter 7), GMES is a key element of the European space strategy drafted by the European Commission and ESA in November 2001.

GMES is a joint endeavour to support Europe's goal regarding sustainable development and global governance, in support of environmental and security policies. The aim is to facilitate the acquisition and distribution of data and information. GMES represents a vital part of Europe's contribution to issues affecting the global environment (see Chapter 9).

ESA develops activities to define the space component of the European GMES Programme, renamed *Kopernikus* on September 16, 2008, including new families of

satellites, called Sentinels, and addressing the issues of standardizing and improving ground segment infrastructures:

- *Sentinel-1 C-band SAR:* sequence of satellites addressing the issue of data continuity for SAR. The immediate priority is to ensure continuity for C-band data.
- *Sentinel-2 Super-spectral:* numerous services of high strategic importance and economic value are being provided based on data from the SPOT and Landsat series of satellites. Sentinel-2 super-spectral will provide continuity and thus guarantee the availability of data to service providers and users.
- *Sentinel-3 Ocean:* this mission will provide an operational basis, in support of oceanographic services that have been developed since 1991 with ERS and ENVISAT. The altimeter part of the mission will contribute to a worldwide operational oceanographic service.
- *Sentinel-4/-5 Atmospheric Chemistry:* elements for atmospheric monitoring, including real-time services related to atmospheric chemistry, pollution, ozone and aerosols. Sentinels-4 and -5 would be space-based systems operating from geostationary (GEO) and low Earth orbit (LEO), respectively.

6.4.4.1 ENVISAT

ENVISAT-1, planned and built with the support of the ESA, was successfully launched from Kourou (French Guyana), March 1, 2002 (see Plate 6.3). This polar orbit advanced satellite for Earth Observation is part of the ambitious and innovative ESA programme, which should ensure continuity in terrestrial surface data collection continuing the activity of ERS-1 and -2 (Box 6.10). ENVISAT with its different scientific and operative sensors supports many research activities regarding water, soil and vegetation, providing measurement of atmospheric parameters, monitoring climatic and environmental changes (Box 6.11).

Box 6.10 Objective of the global and regional mission of the ENVISAT programme

Global mission	Regional mission
Temperature monitoring of the marine surface	Analysis and map of snow and ice
Monitoring of O_3	Coastal pollution monitoring
Monitoring physical atmospheric parameters	Ocean and sea monitoring
(temperature, pressure, $H_2O_{(v)}$, cloud altitude)	Agro-forest monitoring
Pollution monitoring in marine environment	Geologic and mineral resources analysis
Radiative process	Research and application in hydrology
Thermal change at the marine–atmosphere limit	Researches and application in hydrology
Marine productivity of the oceans	Fishiness of the coastal water
Dynamic of the polar calottes	

Box 6.11 Sensors on board ENVISAT, characteristics and main applications

Acronym	Sensor	Characteristics	Relevant applications
MERIS	MEdium Resolution Imaging Spectrometer Instrument	Optical spectrometer 15 spectral bands from 0.412 to 1.05 μm. 300–1200 m geometric resolution	Measure of the reflected radiation. Measure of biophysical of oceanic and coastal water: chlorophyll concentration and suspended material. Atmospheric aerosol measure on marine water. These data are helpful in the comprehension of carbon cycle in the ocean, thermal oceanic regimes and management of oceanic/coastal waters.
ASAR	Advanced Synthetic Aperture Radar	Radar C-band (5.3 GHz ca.) Alternate polarization, Field of view modification capacity Data acquisition in five operative modes geometric resolution from 30 m (Image mode) up to 1 km (Global monitoring mode)	*Land:* forest map, desertification monitoring, topographic measure (InSAR). Geologic application. Agricultural application. Environmental risk: flooding monitoring, hydrocarbon marine pollution monitoring, subsidence and earthquake analysis (Interferometry SAR). *Sea:* characterization of the oceanic surface (waves, tsunami, tides), surface wind monitoring, bathymetry. *Ice:* glacier and ice mapping, iceberg and ship routing monitoring.
AATSR	Advanced Along Track Scanning Radiometer	Radiometer 2 bands visible 1 band near IR 2 bands SWIR 1.6 and 3.7 μm 2 bands TIR 10.7 and 12.0 μm Resolution 1 km Accuracy 0.3 K in SST measurement.	Marine surface temperature measurement (SST), models of oceanic circulation. Analysis of vegetation parameters: biomass, water content, phenology, health. Fire monitoring at global scale. Meteorological anomalies. Atmospheric water vapour analysis and radiative balance.
RA-2	Radar Altimeter	Radar altimeter measuring the echo radar delay from the surface accuracy < 1 ns.	Oceanic topography measurement complementary at the research about the oceanic circulation,

			bathymetry and characteristics of the marine geoid. Wind measurement and of the oceanic waves, meteorological forecasting support. Ice thickness and relationship with the climatic global change. Topographic application.
MWR	MicroWave Radiometer	Microwave radiometer	Analysis of water vapour content in the atmosphere, vertical profiles. Analysis of clouds liquid water content. Research support at the atmospheric energetic balance.. Analysis of soil moisture.
DORIS	Doppler Orbitography and Radiopositioning Integrated by Satellite	Echo Doppler measurer	Precise measure of the orbital position. Researches about glaciers and environmental changes. Topographic applications and studies about the terrestrial gravitational field.
GOMOS	Global Ozone Monitoring by Occultation of Stars	Spectrometer from 0.25 to 0.95 μm). Vertical resolution <1.7 km	Day/night recording. Global ozone budget. Measure of vertical profiles of greenhouse gases: NO_2, NO_3, Cl, F. Water vapour balance in the atmosphere.
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding	Interferometer from 4.15 to 14.6 μm	Independent from solar light. Stratosphere chemistry, spectral measures of atmospheric gases (O_3, H_2O, CH_4, N_2O, and HNO_3). Ozone and CFC monitoring.

Anyway there is a difference in SAR sensor wavelengths (5.33 cm in ENVISAT, compared to 5.12 cm in ERS-1 and 2), which makes the combined use, and thus the interferometric use, of the two systems more difficult, although possible.

The ENVISAT mission involves ESA and several space agencies, as well as the Canadian CSA (*Canadian Space Agency*), for its completion, and it represents one of the most important international space programmes, developed to monitor and understand environmental phenomena at global and regional scale.

The satellite orbits on polar, sun-synchronous path, at 800 km altitude, crossing the equator at 10:00 a.m. (solar time) with a period of 35 days.

The main objectives of the ENVISAT programme deal with the chance to collect environmental data at both global and regional scale; Box 6.10 reports the typologies of collected data for each of the two scales of investigation.

Table 6.4 Swath width acquisition with different incidence angles of the radar system ASAR

ASAR Code	Swath width (km)	Close-range angle (°)	Far-range angle (°)
IS1	108.4 – 109.0	14.1–14.4	22.2–22.3
IS2	107.1 – 107.7	18.4–18.7	26.1–26.2
IS3	83.9 – 84.3	25.6–25.9	31.1–31.3
IS4	90.1 – 90.6	30.6–30.9	36.1–36.2
IS5	65.7 – 66.0	35.5–35.8	39.2–39.4
IS6	72.3 – 72.7	38.8–39.1	42.6–42.8
IS7	57.8 – 58.0	42.2–42.6	45.1–45.3

The *Medium Resolution Imaging Spectrometer Instrument* (MERIS) is an array electronic devices (pushbroom) acquiring data in 15 bands in the visible and near infrared whose width and position can be regulated. The geometric full resolution is 300 m permitting the acquisition of the Earth in 3 days. MERIS full resolution data are collected and processed only under request for regional scale studies: the acquisition is carried out at reduced scale of 1200 m at nadir.

The *Advanced Synthetic Aperture Radar* (ASAR), the largest device on the satellite, was planned to operate at high, medium and low resolution, thanks to different polarization combinations and incident angles. During the processing phase, ASAR data are processed to obtain medium resolution images (150 m), while high-resolution data are processed and made available for regional scale studies under express request by the users.

ASAR can then generate images having a geometric resolution no lower than 30 m and with HH or VV polarization (C-band, 5.331 GHz). The swath width is functional to the incidence angle (15–45°) (Table 6.4). Related thematic fields of application concern land cover and land use, ocean dynamics, glacier movements, deforestation and desertification.

6.4.4.2 ERS-1 and ERS-2 Satellites

ERS-1 (*European Remote Sensing radar*), developed by ESA is the first European polar, sun-synchronous orbit satellite, launched in 1991, followed by the ERS-2 on April 21, 1995 (see Plate 6.3). These two satellites acquire one polarization (VV) and one wavelength (band-C) images; as a consequence, obtained data are limited by technical parameters fixed by the instrument (Box 6.12).

The close acquisition of ERS-1 and ERS-2 images with 1-day displacement (tandem mission), carried out for a few months in 1995–96 period, allowed interesting advances in the knowledge about interferometry and its possible applications. Images were acquired with a period ranging from 3 to 35 days.

The ERS-1 satellite stopped functioning on March 10, 2000 after almost 9 years, three more than planned. In quantitative terms, the two ESA satellites weigh 2516 kg, are 11.8 m long and the SAR antenna measures 10 m × 1 m.

The devices installed on the ERS-2 satellite are identical as on ERS-1, together with the sensor *Global Ozone Monitoring Experiment* (GOME) not present on the

first satellite. Both the satellites carry *Active Microwave Imager* (AMI) and *Radar Altimeter* (RA), as well as other instruments for atmospheric studies and orbital monitoring.

Box 6.12 Instruments on board the ERS-1 and ERS-2 satellites

Acronym	Sensor	Relevant characteristics
AMI	Active Microwave Instrument	In band-C (5.3 GHz) operating as SAR (Synthetic Aperture Radar) antenna and as scatterometer for measuring wind speed. By means of the SAR antenna 2D images can be recorded with 26 m geometric resolution in transversal direction orbit and at azimuth with resolution from 6 to 30 m. The position of the antenna is parallel to the orbit obtaining images with swath width < 100 km. The scatterometer collects oceanic surface wind speed data used in models and climatological database. The wind speed modifies the roughness of the marine surface inducing modification of the backscattered signal. Geometric resolution >4.5 km and swath width 500 km
RA	Radar Altimeter	in band KU with 13.8 GHz (λ: 2.6 cm). Designed for measuring the returning time of the radar signal, vertically transmitted, from the oceanic and glacial surfaces. Data collection on wave high, wind speed (combined with the AMI scatterometer) and tides. The instrument is appropriate for comprehension of oceanic dynamics at global scale and continuous glacier monitoring of the Earth's surface.
ATSR	Along-Track Scanning Radiometer	Earth imagery in the near infrared with 1 km geometric resolution. ATSR-1, on board ERS-1 is without the visible band of ATSR-2, on board ERS-2. The instrument is appropriate for the monitoring of the atmosphere humidity, and its evolution ATSR-2, for analysis of vegetation condition.
GOME	Global Ozone Monitoring Experiment	On board ERS-2. Absorbing spectrometer operating in the UV and VIS ranges for monitoring of ozone, gas traces and aerosol in the stratosphere and troposphere.
MS	Microwave Sounder	Gives data concerning the humidity content in the atmosphere
PRARE	Precise Range and Range Rate Equipment	Defines the altitude and the satellite track
LRR	Laser Reflector	Defines the satellite altitude of the orbit and registers the 6.5 Gbit of data acquired in one orbit.

6.4.5 ASI (Italy) Space Programme

6.4.5.1 COSMO/SkyMed

COSMO/SkyMed (*Constellation of Small satellites for Mediterranean basin Observation*) programme was conceived as an operative satellite system for EO addressed to marine, coastal, terrestrial resource management and environment and disas-

Table 6.5 Operative characteristics of the constellation COSMO/SkyMed

Parameter	Unit	Value
Satellite	Number	4 radar + 2 optical
Orbit altitude	km	694.80
Speed	km/s	7.5
Type of orbit	–	Sun-synchronous, near polar
Inclination	degree, °	98° 12'
Orbit period	minutes	99
orbits/day	number	14
Incidence angle	degree, °	25–60 (radar)
Acquisition	–	nadir and off-nadir 30° (optical)
Acquisition limits	degree, °	± 81
Type of instrument	–	optical/electronics
Sensor	–	CCD
Active radar band	cm	3 band X
Passive optical bands	Number	1 Pan + 4 multispectral
Interferometry	–	Yes (radar)
Stereoscopy	–	Yes (optical)

ter monitoring, mainly centred in the Mediterranean area, also able to respond to requests from other regions worldwide (Table 6.5).

The project is based on the realization of a *space segment* constituted by a constellation of low orbit small satellites, equipped with both optical and radar instruments, and its *ground segment* for monitoring, planning, receiving and processing observed data, and on simultaneous access by different users also through the use of telecommunication satellites.

This system, called ORFEO (*Optical and Radar Federated Earth Observation*), is collaboration between France and Italy on a dual high-resolution Earth observation system (Fig. 6.24).

The constellation comprises six satellites, four equipped with a Synthetic Aperture Radar (SAR) in X-band (3 cm) microwave region of the spectrum, and two identical optical instruments, provided with high-resolution multispectral (MS) and panchromatic (Pan) devices able to operate in different acquisition modes.

The SAR-X satellites COSMO/SkyMed 1 and 2 were launched respectively on June 7 and December 10, 2007, COSMO/SkyMed 3 on October 25, 2008, from the Space Centre in Vandenberg,California, by the rocket Boeing Delta II.

The project is characterized by:

- quick availability (up to the final user) in 24 h after the acquisition;
- proper geometric resolution of the images, with better detail definition at the requested scale of analysis;
- any-time acquisition capacity both night and day (radar);
- single passing synoptic view;
- along-track stereoscopic acquisition in the single passing;
- immediate interpretation of the data, quick distribution and global accessibility.

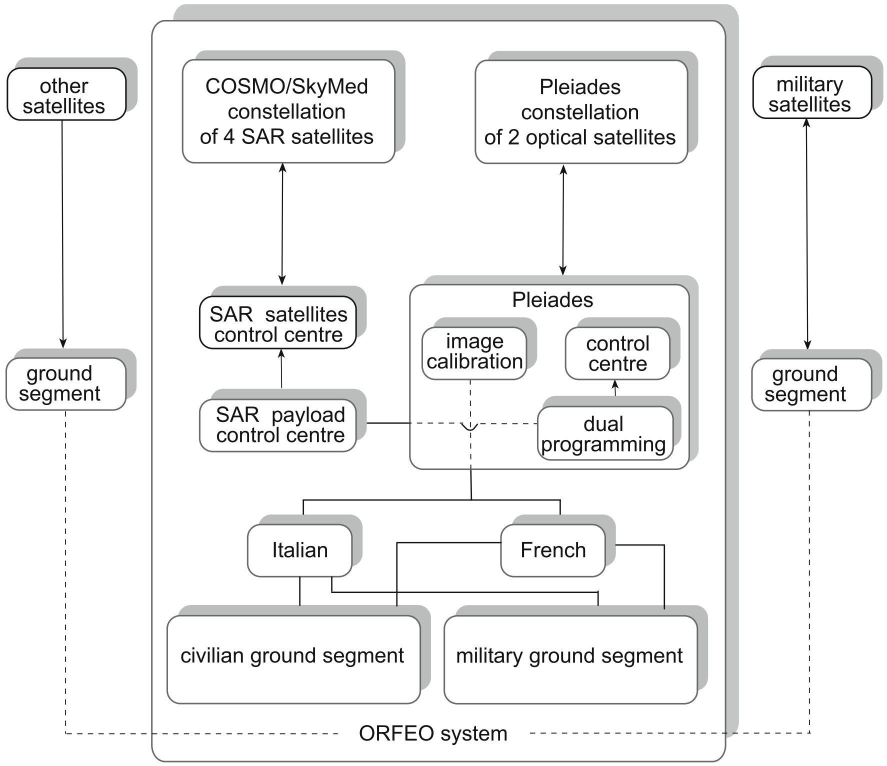


Fig. 6.24 ORFEO system constituted by a constellation of four SAR (COSMO/SkyMed) and two optics (Pleiades) low-orbit small satellites, a result of Italian/French cooperation

The choice to combine an *optical system* (daylight acquisition without clouds), made by a sub-constellation of two sun-synchronous midnight–midday orbit satellites, equipped with a *radar system* (anytime), and a sub-constellation including four sun-synchronous dawn-sunset orbit satellites, having 5-day period, satisfies the request of anytime capacity and high temporal resolution.

The sensors can acquire in different modes, in terms of both swathes width and spatial resolution. The capacity to record large areas through one passing using a mosaic technique, maintaining high geometric resolution, constitutes one of the main characteristics. The optical system can acquire stereoscopic pairs and tri-stereo in several configurations (350 km × 20 km or 150 km × 40 km). The swath ground width is about 20 km.

The geometric resolution of panchromatic images (Pan) is 0.7 m at nadir, in any case less than 1 m for acquisitions with an angle lower than 30°. Multispectral images (MS) geometric resolution at nadir is 2.80 m.

The main characteristics of COSMO/SkyMed Programme are reported in Tables 6.6 and 6.7.

Table 6.6 Operative characteristics of the radar constellation in band X of the 4-satellites COSMO/SkyMed

SAR	Unit	High	Resolution medium	Low
Swath width	km	24–42	> 120	200
Geometric resolution	m	<3	25	100
Band	–	X	X	X
Polarization	–	HH	HH	HH
Incidence angle	degree, °	25–60	25–60	25–60

Table 6.7 Operative characteristics of the optical constellation of two-satellite, high-resolution camera, on board COSMO/SkyMed

High-resolution camera	Unit	Value
Swath width	km	15
	km	> 50
Geometric resolution	m	0.7 panchromatic (Pan)
(nadir)	m	2.8 multispectral (XS)
Type of acquisition	–	stereo, mosaic, swath
Spectral range Pan	μm	0.45–0.85
XS	μm	0.43–0.55 blue
		0.49–0.61 green
		0.60–0.72 red
		0.75–0.95 near infrared
View angle	Degree, °	± 35

The optical part *Pleiades* is described in the *French Space Agency* (CNES) programme.

6.4.6 CNES (France) Space Programme

The French space agency had and still has a relevant role in developing projects for EO through satellite platforms.

CNES works actively to implement the space component of GMES through the ESA and French Earth Observation Programme developing since the late 1980s the SPOT series, the *Jason* ocean observatory and the *Pleiades* optical systems.

In collaboration with the ERS European programme, CNES provides support to satellite platforms similar to the SPOT series, developing the AATSR radiometer (six bands in the visible and infrared ranges, Box 6.11) and maintaining the efficiency of the archive and data processing Centre (CERSAT).

Within the ENVISAT project, the CNES (*Centre National d'Etudes Spatiales*) developed the device DORIS (*Doppler Orbitography and Radiopositioning Integrate by Satellite*), already installed on SPOT 4, a very important device in maintaining the orbit and collecting topographic data.

6.4.6.1 SPOT Satellites

SPOT 1, the first satellite of the SPOT series (*Système Pour l'Observation de la Terre*), planned and realized mainly by the CNES, has been operative since February 22, 1986, followed in 1990 and 1993 by SPOT 2 and 3 which have characteristics similar to the previous. The satellite SPOT 2 lost its recording capability in 1993 but can still send direct images, while SPOT 3 has had interrupted operativity since 1997.

Within the European programme for the global monitoring of vegetation, called *Vegetation Programme*, supported by France, European Commission, Belgium, Italy and Sweden, the satellites SPOT 4 and 5 were planned and developed integrating and improving the characteristics of the previous. The aim of this project is to monitor parameters of the terrestrial surface, agriculture, farming and forest production and of the biosphere up to a daily frequency over the whole of the Earth's surface also extending medium geometric resolution to the medium infrared.

SPOT 4 was launched in 1998, with an updated version of the *High-Resolution Visible* (HRV) device, the *High-Resolution Visible and Infrared* (HRVIR), having one more band in the medium infrared (1.58–1.75 μm), and the *Vegetation* instrument (VGT), developed within a European programme for vegetation global monitoring. SPOT 5 carries an even more improved instrument, the *High-Resolution Geometry* (HRG), using the same spectral bands as the HIRVR but with 10 m geometric resolution in the multispectral and 3–5 m resolution in the panchromatic (Table 6.8).

Table 6.8 SPOT satellites series: instruments, spectral bands and geometric resolutions

Satellite	Sensor	Spectral bands (μm)	Resolution (m)
SPOT 1, 2	HRV- P	0.51–0.73 panchromatic	10
and 3	HRV -XS	0.50–0.59 green	20
		0.61–0.68 red	
		0.79–0.89 near infrared	
SPOT 4	HRVIR - P	0.51–0.73 panchromatic	10
	HRVIR -XS	0.50–0.59 green	20
		0.61–0.68 red	
		0.78–0.89 near infrared	
		1.58–1.75 medium infrared	
SPOT 4 and 5	Vegetation	0.43–0.47 blue	1000
		0.61–0.68 red	
		0.78–0.89 near infrared	
		1.58–1.75 medium infrared	
SPOT 5	HRG - P	0.51–0.73 panchromatic	5 and 3
	HRG - XS	0.50–0.59 green	10
		0.61–0.68 red	
		0.78–0.89 near infrared	
		1.58–1.75 medium infrared	

Operative SPOT satellites have circular near-polar, sun-synchronous orbit at 832 km altitude. They pass over the same area, at the same time (10:30 a.m. local time at the equator) every 26 days (nadir orientation).

Optical acquisition systems are based on scanner with no moving devices (push-broom) (Box 6.13).

Box 6.13 SPOT satellites series.

Satellite	Sensor	Launch date	End operativeness
SPOT 1	HRV	February 22, 1986	January 5, 1999
SPOT 2	HRV	January 22, 1990	Operational
SPOT 3	HRV	September 26, 1993	November 14, 1997
SPOT 4	HRVIR and Vegetation	March 24, 1998	Operational
SPOT 5	HRG	May 4, 2002	Operational

The panchromatic (Pan) array has 6000 sensors positioned in a linear array, with 10 m × 10 m resolution, and the multispectral HRV and HRVIR array has 3000 sensors in line, with 20 m × 20 m geometric resolution. In SPOT the two instruments′ view axis can be oriented vertically (nadir vision) as well as with a view angle of ± 27° from the vertical, with 0.6° rotation intervals. In the case of nadir vision, the scanning width is 117 km (60 + 60 – 3, where 60 km is the width of the acquisition surface from each instrument and 3 km is the overlapping of the two swathes); in the case of the slant vision of 27°, the scanning line is 950 km width. Each sensor oriented at ± 27° from the vertical covers a ground swath of 475 km, generating considerable pixel warping.

Thanks to this characteristic the same area can be observed with a period of 1–4 days, using different acquisition angles and obtaining stereoscopic images with short intervals.

Two identical *High-Resolution Visible* (HRV) are the operative instruments on board the first three satellites of SPOT series, planned to operate in two modes both in the visible and in the infrared:

- *multispectral* in three bands, specifically pointed to the spectral response of vegetation, with 20 m × 20 m spectral resolution:
  - two of the bands are in the visible, 0.50–0.59 μm, 0.61–0.68 μm;
  - one is in the near IR, 0.79–0.89 μm;
- *panchromatic*, 0.51–0.73 μm, with 10 m × 10 m geometric resolution.

Vegetation was planned for daily multi-temporal monitoring of crops and natural vegetation at global and regional scale and the fields of application include environmental studies and biomass productivity forecasting. The field of view (FOV) is 2250 km. The adopted optical–electronic acquisition system, made by 2259

elements, maintains 1 km geometric resolution constant over the FOV also increasing pixel warping nearer the borders of the observed areas. HRVIR and *Vegetation* acquisition are simultaneous with coherent geometrical references; this facilitates sampling operations and multi-resolution studies taking advantage of both *Vegetation* multi-temporality and HRVIR high-spectral resolution. The red band, absorption peak, and near infrared band, reflection peak of vegetation, are comparable with the NOAA-AVHRR bands. Their ratio, normalized difference and other equivalent algorithms are used for studying biomass development. *Vegetation* misses bands in the emitted infrared for the study of bodies' temperature.

The archiving of data from SPOT 4 is managed by Kiruna−Esrange receiving station, in northern Sweden, to which the satellite transmits the data stored in the memory on board (3–4 times/day).

Data from *Vegetation*, collected in Kiruna, are processed in Mol Centre, Belgium, and made available after 2 days since the acquisition.

6.4.6.2 Jason

The satellite Jason-1 for operational oceanography took over in December 2001, following the experimental TOPEX/Poseidon programme, launched in 1992. Focal objectives of the mission include scientific research and services such as weather bulletins, charts to aid navigation and real-time sea monitoring.

Jason-2, the first of several, took over June 20, 2008 on the same orbit as Jason-1. It has been planned for studying and observing oceans and climate, by the *Ocean Surface Topography Mission* (OSTM) to cover the entire planet.

OSTM collects and distributes high-resolution data on ocean surface topography, which reveal the speed and direction of ocean currents, sea surface height and measurement of the Sun's energy stored by the ocean with operational continuity.

The radar altimeter on Jason-1 and -2 modulates signals reflected from the surface of the ocean. The signal is used to calculate the distance satellite−sea surface, with accuracy within 2 cm.

6.4.6.3 Pleiade

Since 1997 CNES is studying the use of smaller satellites (medium spacecraft size of about 900 kg instead of 3 tons as for the SPOT-5), resulting in the 3S platform concept (*Small Satellite System*) standing for *Suite de Système SPOT* or for *SPOT Successor System*. The objectives are cost reduction, technological innovation, user services and performance upgrades for a new generation of optical imaging satellites, referred to as *Pleiades*.

The Pleiades optical system is being developed under CNES supervision as part of a joint effort with Italy, alongside the COSMO/SkyMed radar imaging system, under the responsibility of the Italian Space Agency (ASI).

The French–Italian cooperation gives large and small-area optical and radar imagery at high and low resolutions.

With two satellites operating simultaneously, Pleiades will be able to acquire imagery of any point on the globe, every day. The optical Pleiades provide imagery with resolution of 0.70 m. The superior agility of the Pleiades high-resolution satellites will afford the ability to acquire optical imagery in three distinct modes:

- *single-pass:* the satellite is able to acquire a mosaic of images covering several adjacent ground swaths (120 km wide by 110 km length) on each satellite pass;
- *tri-stereo:* the satellite's ability to lean fore and aft along its orbital track gives the opportunity to image the same area from different angles on the same pass; this mode will generate near-nadir images and relief perspectives from associated stereoscopic views;
- *multi-spot:* the ability to acquire multiple views of an area of interest the same day, fundamental in case of crisis situations.

6.4.7 FSA (Russia) Space Programme

The Federal Space Programme of the *Rosaviakosmos, Russian Space Agency* (RSA, formerly RKA), which in 2004 was restructured as the *Federal Space Agency* (FSA) with the name of *Roskosmos,* in addition to further development for the GLONASS navigation system, includes an intense development of EO satellite systems.

One of main objectives of the programme is to develop and to maintain the constellation of meteorological satellites (three polar-orbiting, one of them oceanographic and two geostationary satellites).

The environmental satellites include two series:

- Kanopus-V, for earthquake detecting and monitoring as well as for remote sounding of the atmosphere;
- Resurs-P satellite series to be developed to provide detailed Earth surface observations.

The Federal Space Programme also covers the support and development of the Russian segment of the *International Space Station* (ISS), the building of advanced Angara and Soyuz-2 launchers, support at the cosmodrome, ground control and tracking infrastructure.

In 1998 the *Russian Space Agency* joined the ISS Programme, steering Russia into becoming a fully fledged partner and holding to its commitments in relation with other agencies such as ESA, NASA, JAXA and the Canadian Space Agency.

6.4.7.1 MIR Space Station

Russian space activities were represented for 15 years by the orbiting station MIR (meaning *Peace* and *World*), launched on February 10, 1986 completing more than 85,000 orbits around the Earth. The MIR project is still the most relevant space

project developed by FSA and heritage of the space programmes of the former Soviet Union.

Several modules were installed during the MIR operativity in order to collect data about multidisciplinary scientific topics. On March 23, 2001, the destruction of the 136 ton Russian space platform was driven from the ground control station, as it was thought too expensive and obsolete, making it going back to the atmosphere and then destroying and sinking it into the Pacific Ocean.

The core of the research in environmental topics was the PRIRODA module (*nature*, in Russian), on which several sensors were set providing ecological, oceanographic and climatological data.

6.4.7.2 Priroda

PRIRODA module was planned in 1993 and launched in April 1996. The relevance of this programme is the combined use of active and passive instruments in several spectral bands. The multi-sensor approach to the issue of environmental monitoring has been complex but could provide a large amount of information.

Also German and French agencies contributed to the programme.

The operative tools on board PRIRODA, conceptually still valid, are

- *Passive microwaves*

IKAR-N: *nadir microwave radiometer system* operates at centimetric wavelengths (0.3, 0.8, 1.35, 2.25, 6.0 cm), with resolution up to 60 km.

IKAR-P: *scanning microwave radiometer system* generates 750 km swathes and resolution up to 75 km, with wavelength of 2.25 and 6 cm.

IKAR-DELTA: has a radiometer sensing in four wavelengths (0.8, 1.35, 2.25 and 4 cm) able to generate 400 km swath with variable geometric resolution ranging from 8 to 50 km.

Active Microwaves

SAR TRAVERS: operating at 9.3 and 23 cm wavelengths generating swath width of 50 km and with pixel ground resolution of 150 m.

Optical

MSU-SK: is a multispectral scanner having the same bands as Landsat MSS. It has two bands in the visible (0.5–0.6 μm and 0.6–0.7 μm, with 80 m geometric resolution) and two bands in the thermal infrared, 10.3–11.8 μm, with 300 m resolution.

MSU-E2: its 10 m resolution at nadir obtains information about terrestrial surface at local scale. It is characterized by two bands in the visible (0.5–0.6 μm and 0.6–0.7 μm) and one in the near infrared (0.8–0.9 μm). Swath width is 2 km × 24 km.

MOS-A: provides data about O_2 absorption in the near infrared, vegetation and soil conditions indices. It operates in four bands with interval in the near infrared,

0.755–0.768 μm. The geometric resolution is 2.87 km × 2.87 km and the swath width is 80.5 km.

MOS-B: this sensor provides the same typology of data as MOS-A; it operates in 13 bands in the visible/near infrared from 0.408 to 1.01 μm, with 0.7 km × 0.65 km resolution and swath width of 82 km.

ALISA (LIDAR): active sensor working in the visible wavelength of 0.532 μm with 150 m geometric resolution.

MOMS-2P (MOMS-02 on PRIRODA): the *Modular Opto-electronic Multispectral Stereoscanner* device entirely constructed by German technology from DLR Space Agency is based on modern electronic technology that provides narrow spectral bands with an acceptable signal-to-noise ratio. Sensors capability of stereo acquisition of three scanning along the flight path (± 21° and nadir) assured the realization of Digital Elevation Models (DEM) from data recorded in the same irradiance conditions.

6.4.7.3 Resurs-01

The third and fourth satellites of this series (RESURS-01 n.3 and n.4) were launched respectively in November 1994 and July 1998. They had on board sensors not new in conceptual-planning terms, but useful to partially solve the problem of the size of the area of acquisition. In fact these images have geometric resolution of 170 m, between AVHRR 1.1 km and *Thematic Mapper* 30 m, detecting 600 km × 600 km scenes: more than nine Landsat scenes in one Resurs-01 (Fig. 6.25).

The instrument has the acronym MSU-SK (with different characteristics than MSU-SK on board PRIRODA) and is a multispectral scanner with the same bands as Landsat MSS, with 50 m ground resolution and 100 km swath width. The sensor MSU-SK also has a band in the thermal infrared with 600 m resolution in the spectral range 10.4–12.6 μm, comparable with NOAA-AVHRR.

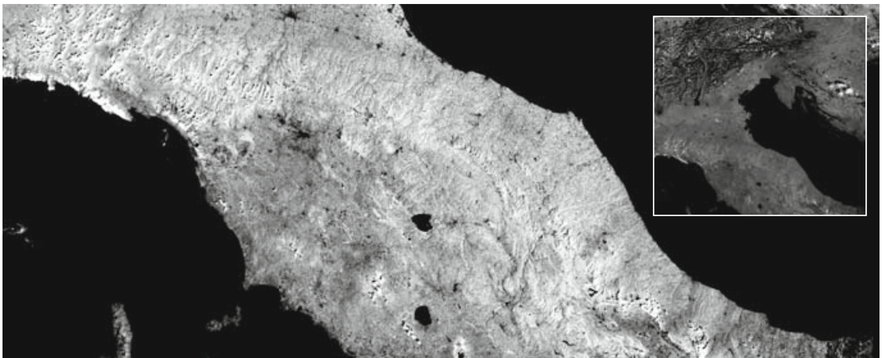


Fig. 6.25 Near infrared band of the sensor MSU-SK of the satellite Resurs-01, Central Italy. The geometric resolution is of 160 m. In the top right, the full scenes 600 km × 600 km

Resurs-01 frequency of acquisition is 4 days, with free cloud cover.

The combination of the synoptic view and close frequency with the medium ground resolution make these data suitable for dynamic land monitoring and production of land use/land cover maps at scales between 1:250,000 and 1:1,000,000.

6.4.7.4 Photographic Cameras on Board Satellites

Opportunities about the availability of EO data increased after 1989 subsequent to the deep changes which occurred in Russia: photographic acquisitions from space became available for civilian purposes and not only for military use. The photographic version was later reproduced also in digital format, in order to allow a better diffusion of data. The acquisitions are realized by metric cameras using films also 3 km long, on stations orbiting at altitude from 270 to 470 km, with or without crew.

The interest in this data is due to the better ground resolution ranging from 2 to 10 m. The spectral range of these photographic products can be both panchromatic and multispectral; the temporal interval of acquisition is very discontinuous, and varies from 14 days to some months.

These data are collected by three different photographic cameras having the following main characteristics:

- KFA-1000: acquisitions from 1984 to 1993; 20 km $\times$ 120 km scenes with 5 m geometric resolution in panchromatic mode and 10 m in multispectral. Two are the used spectral regions, in the visible: 0.57–0.68 μm, and in the near infrared: 0.68–0.81 μm;
- KFA-3000: acquires photos having 3 m resolution. It works in the interval 0.6–0.7 μm. The swath width corresponds to 1% of the height;
- KATE 200: acquires an area of 225 km $\times$ 235 km with geometric resolution of 20 m. The used spectral regions are three, two in the visible: 0.50–0.60 μm, 0.60–0.70 μm, and one in the near infrared: 0.70–0.85 μm;
- MK4: acquisitions from 1998 to 1995; it covers 170 km $\times$ 170 km with 10 m geometric resolution. The used spectral regions are five, in the visible: 0.43–0.46, 0–46–0.51, 0.51–0.56, 0.63–0.69 μm, and near infrared: 0.81–0.90 μm.

COSMOS series satellites, born for military use, known since 1988 with the acronym SPIN-2, were equipped with two metric panchromatic photographic cameras, TK-350 (acquisitions from 1982 to 1998), covering an area of 200 $\times$ 300 km^2 with 80% overlapping on the two swaths, obtaining the stereoscopic vision, and KVR-1000 (acquisitions from 1985 to 1992) with 2 m resolution (Table 6.9). Photographs are then scanned at high resolution and recorded on digital support. Although automatic positioning systems and laser systems verifying the heights are supposed to be on board, SPIN-2 images ground geometric precision is about $\pm$ 20 m.

However SPIN-2 images have anticipated the possible applications available from commercial satellites, having resolution lower than 1 and 2 m ground geometric accuracy. At the moment interesting data are collected by the *Space Mapping*

Table 6.9 Parameters of the acquisition devices TK-350 and KVR-1000 of the Russian system KOMETA

Parameter	Unit	Topographic Camera TK-350	High-resolution Camera KVR-1000
Focus	mm	350	1000
Resolution	m	10	2
Average scale	–	1: 660,000	1: 220,000
Photo size	cm	30 × 45	18 × 72
Scene size	km	200 × 300	40 × 160
Stereoscopic overlap	%	20–40–60–80	NO

System KOMETA. This system was planned with the purpose of producing stereographic topographic photographs and high-resolution panoramic photo-images in order to produce topographic maps at 1:50,000 scale and smaller, with advantages where it is not possible to collect Ground Control Points. The system externally uses the positioning data from the *Global Navigation System* (GLONASS) and internally the satellite has a laser altimeter and Doppler tools for correctly positioning the points.

The main characteristic of the system is to allow the production of Digital Elevation Model (DEM) and to have 2 m geometric ground resolution. The archive of high-resolution images contains data acquired from 1981.

The products may have the following characteristics:

- georeferenced images with 20 m planimetric error, without Ground Control Points, and DEM with 10 m precision in height;
- photographic products in the formats: positive film, negative film, contact prints on paper, photographic enlargements (TK-350 from 0.7× up to 16× (1:14,000) without quality loss; TK-1000 from 0.7× to 22× (1:10,000) without quality loss);
- digital copies with high-precision (density) scanning;
- TK-350: 14 μ corresponding to 10 m on the ground;
- KVR-1000: 7 μ corresponding to 2 m on the ground;
- digital orthophoto (Root Mean Square RMS ± 20 m, ground resolution: 2 m);
- digital orthophoto with Ground Control Points (RMS ± 1–2 image pixel in planimetry);
- Digital Elevation Model (DEM) (RMS ± 20 m in planimetry; RMS ± 3–5 m in height);
- Digital Terrain Model (DTM) using 1500–2000 points per km^2 (TK-350) and 150–200 points per km^2 in KRV 1000 (RMS ± 20 m in planimetry; RMS ± 5 m in height).

6.4.7.5 Resurs-DK

Partial imagery with 1 m geometric resolution (DK-1) has been available since 1992. These data are acquired by remote sensing panchromatic systems. The acquisition camera has a wide swath that reduces the complex mosaic process.

Table 6.10 The main technical characteristics of the system Resurs-DK

Linear resolution	Unit	Sensor
Panchromatic	1 m	DK-1
Multispectral	2–3 m	DK-2
Spectral bands		
Panchromatic	0.58–0.8 μm	DK-1
multispectral	0.5–0.6 μm	DK-2
	0.6–0.7 μm	DK-2
	0.7–0.8 μm	DK-2
Track and orbit		
Swath width	28.3 km	DK 1-2
Swath length	448 km	DK 1-2
Orbit altitude	350 km	DK 1-2

The products, with resolution of 1 m in digital orthorectified format, are images as well as elaborated products like high-resolution digital maps and GIS thematic layers. The accuracy of these products can reach the definition of equivalent maps at 1:10,000 and 1:5000 scales.

Similarly the Resurs-DK2 acquires with 2 m linear geometric resolution.

The system Resurs-DK2 is an opto-electronic system original conception drawn in order to be compatible with data receiving stations able to receive via radio in the international frequency of 8.2–8.4 GHz (Table 6.10).

It is possible to generate a personalized acquisition plan with 1 m resolution (DK1) over areas where a better detail is required and 2 m resolution (DK2) over the others (Fig. 6.26).

Fig. 6.26 DK2 images of the historical centre of Rome: Saint Peter's and the Vatican

The imagery can be used for the following purposes:

- topographic and thematic maps;
- digital orthophotos;
- Digital Terrain Model (DTM) and Digital Elevation Model (DEM);
- digital image processing for analysis in agriculture and forestry, vegetation, environmental monitoring, hydrology, soil maps, geological survey, photo-maps, cartography;
- development of thematic layers including urban areas, road network maps, hydrographic network;
- functional analysis of urban areas;
- scenarios about current land use conditions;
- landscape change and geomorphological situations analysis.

6.4.8 ISRO (India) Space Programme

The prime objective of the *Indian Space Research Organisation* (ISRO) is to develop space technology and its application. ISRO has established two major space systems, *Indian National Satellite System* (INSAT) for communication, television broadcasting and meteorological services, and *Indian Remote Sensing Satellites* (IRS) sun-synchronous satellites for natural resources′ monitoring and management (Plate 6.3). Furthermore, the launcher vehicles, *Polar Satellite Launch Vehicle* (PSLV) and *Geosynchronous Satellite Launch Vehicle* (GSLV), to place INSAT and IRS satellites in the required orbits, have been developed.

The first satellite, IRS-1A, was launched in 1988, followed by IRS-1B, in 1991. Both had two sensors, LISS-1 and LISS-2 (*Linear Imaging Self-scanning Sensor*), operating in four spectral bands in the blue, green, red and near infrared regions, with a ground resolution of 72.5 m for LISS-1 and 36.25 m for LISS-2, and 22 day period. Without recording systems on board, data could be collected only in the domain of the habilitated receiving stations.

In December 1995, the satellite IRS-1C was launched, followed by the IRS-P3 and IRS-1D, on which a recording device was placed collecting and transmitting data to the several ground stations habilitated. IRS-1C sensors' characteristics are reported in Box 6.14.

The digitalization at 7 bit of LISS bands limits spatial data and radiometric resolutions, reducing the capacity of class's separation in classifications.

IRS-1D satellite, launched on September 29, 1997, entered elliptic orbit instead of circular after separation from the rocket. For this reason, although the sensors on board are identical to the ones on the IRS-1C, the geometric resolution and the swath width are functions of the distance from the Earth. In Table 6.11 each instrument resolution from the perigee and apogee of the satellite is reported.

Box 6.14 IRS-1C satellite sensors

Acronym	Sensor	Characteristics	Relevant applications
PAN	Panchromatic camera	5.8 m resolution, 0.50–0.75 μm single band, swath length 70 km (nadir) and 90 km	Geological and geomorphological studies, urban planning, stereoscopy and generation of Digital Terrain Model (DTM)
LISS-3	Linear imaging Self-scanning sensor	0.52–0.59 μm green 0.62–0.68 μm red 0.77–0.86 μm near IR geometric resolution 23 m swath width 141 km 1.55–1.70 μm medium IR geometric resolution 69 m swath width 148 km Dynamic range 7 bit	Vegetation typology and vegetation health, soil moisture Cycle of 24 days
WiFS	Wide Field Sensor	0.62–0.68 μm visible 77–0.86 μm near IR swath width 810 km geometric resolution 188 m	Vegetation study Cycle of 5 days

IRS-P3 carries the WiFS having the same bands of IRS-1D plus one in the interval 1.55–1.75 μm, with 188 m ground resolution. The three WiFS bands have spectral compatibility with bands 3, 4 and 5 of Landsat TM and ETM+. The combination of sensors on IRS-1C, 1D and IRS-P3 gives a temporal resolution less than 2 days (Table 6.12).

IRS-P4 (*OceanSat*) has a payload with *Ocean Colour Monitor* (OCM) and *Multi-frequency Scanning Microwave Radiometer* (MSMR) and *Radio Occultation for Sounding the Atmosphere* (ROSA) instruments.

The satellite IRS-P6 (*ResourceSat-1*) was launched on October 17, 2003, in sun-synchronous orbit, at 817 km altitude, with three sensors on board:

- *Linear Imaging Self-Scanner-4* (LISS-4) with 6 m resolution and 25 km swath width;
- *Linear Imaging Self-Scanner-3* (LISS-3);
- *Advanced Wide Field Sensor* (AWiFS) acquisition with 80 m resolution and 1400 km swath.

Box 6.15 Characteristics of the IRS satellite series

Table 6.11 Geometrical resolutions of the instruments on board the satellite IRS 1D at perigee and apogee

IRS 1D instruments	Unit	Apogee	Perigee
Pan	m	5.8	5.2
LISS-III	m	23.5	21.2
WiFS	m	812	728

Table 6.12 Operative characteristics of the satellite-sensor systems NOAA, Resours-01, Landsat, SPOT, IRS

Parameter	NOAA-11	Resurs-01	Landsat 4,5	Landsat 1,2,3	SPOT		IRS-1C and 1D		IRS-P3
Temporal Res. (g)	0.5	4	16	18	1–26			24	
Orbit altitude (km)	833/873	678	705	920	832			817	
Radiometric resolution	10 bit	6	8 bit	7 and 6 bit	6 bit	8 bit	6	7	7
Scene size (km2)	2.700 × 2.700	600 × 600	185 × 185	185 × 185	60 × 60	60 × 60	70 × 70	141 × 141	810 × 810
MegaByte per band	9	12.2	40	5.3	36	9	196	37.2	18.5
Sensor	AVHRR	MSU-SK	TM	MSS (L 1-5)	PAN	HRV	PAN	LISS-III	WiFS
Spectral res. [μm]									
Panchromatic					0.51–0.73		0.50–0.75		
Blue			0.45–0.52						
Green		0.50–0.60	0.52– 0.60	0.50–0.60		0.50–0.59		0.52–0.59	
Red	0.58–0.68	0.60–0.70	0.63–0.69	0.60–0.70		0.61–0.68		0.62–0.68	0.62–0.68
Near IR		0.70–0.80		0.70–0.80				0.77–0.86	0.77–0.86
Near IR	0.72–1.10	0.80–1.10	0.76–0.90	0.80–1.10		0.79–0.89			
Medium IR			1.55–1.75					1.55–1.70#	1.55–1.75
Medium IR	3.55–3.93								
IR Thermal	10.3–11.3	10.4–12.5*	10.4–12.5*						
IR Thermal	11.5–12.5								
Medium IR			2.08–2.35						
. Geometric res.	1100	170 600*	30 120*	80	10		5.8	23 69#	188
Pixel / line	3000	3500 1000*	6200	2300	6000	3000	12,068	6100 2050#	4300
Pixel per km2	0.82	34.6 2.8	1.111	156	10,000	2500	29,726	1890–210#	28.3

Box 6.15 Characteristics of the IRS satellite series

Satellite	Acronym	Sensor	Launch date and characteristics
IRS-1A and 1B	LISS-1 LISS-2	Linear Imaging Self-Scanner1 and 2	1988 and August 1991
IRS-1C and 1D	Pan LISS-3 WiFS	Panchromatic Linear Imaging Self Scanner 3 Wide Field Sensor	December 1995 and September 29, 1997
IRS-P3	OCM WiFS	Ocean Colour Monitor Wide Field Sensor	March 21, 1996
IRS-P4	OCM MSMR	Ocean Colour Monitor Multi-frequency Scanning Microwave Radiometer	June 3, 1999
IRS-P5	VHRP	Very High-Resolution Panchromatic camera	Planned for cartographic applications
IRS-P6	LISS-3 LISS-4 AWiFS	Linear Imaging Self-Scanner 3 and 4 Advanced Wide Field Sensor	October 17, 2003 Applications in agriculture.

Its applications are focused on agricultural monitoring.

ISRO also launched a satellite for cartographic applications, *CartSat-1* with a *Very High-Resolution Panchromatic Camera* (VHRPC) on board. It carries two Panchromatic (PAN) cameras for black-and-white stereoscopic acquisitions in the visible region of the electromagnetic spectrum. The swath covered by the PAN cameras is 30 km and their spatial resolution is 2.5 m. The cameras are positioned on the satellite to collect, as an option, near simultaneous imaging of the same area from two different angles.

6.4.9 JAXA (Japan) Space Programme

The Japanese Space Agency NASDA (*National Space Development Agency of Japan*) merged on October 1, 2003 with the other two Japanese institutions, ISAS (*Institute of Space and Astronautical Science*) and NAL (*National Aerospace Laboratory of Japan*) to create a new agency: JAXA (*Japan Aerospace eXploration Agency*). This fusion aimed to give a more systematic approach to space exploration, which requires a wide knowledge, from the basic research to the development of innovative sensors up to practical applications of the employed technology (Box 6.15).

Several projects were developed by the JAXA since the 1980s. The first two Japanese satellites for sea observation collecting data about environmental resources, respectively MOS-1 and MOS-1b, were launched in the space in 1987 and 1990; MOS-1 stopped its operations in 1995 and MOS-1b in 1996. Three instruments constituted the payload:

- MESSR (*Multispectral Electronic Self-Scanning Radiometer*), scanning electronic radiometer made by a double-camera system parallel to the flight direction;
- VTIR (*Visible and Thermal Infrared Radiometer*), radiometer with scanning direction perpendicular to the direction of the orbit;
- MSR (*Microwave Scanning Radiometer*), scanning radiometer acquiring in the radio waves range.

On February 11, 1992 the satellite JERS-1 (*Japanese Earth Resources Satellite*) was launched from the Japanese station of Tanegashima (Kagoshima), into a sun-synchronous orbit at 568 km altitude, with the purpose of observing the Earth's surface for the systematic monitoring of environmental resources through optical sensors and radar. JERS-1 mission ended on October 12, 1998.

The active microwave SAR system operated in L-band (23 cm) with HH polarization. The geometric resolution was 18 m $\times$ 24 m, with 75 km swath, 35° depression angle, 44-day period. One interesting issue was the complementarity with ERS-1 and ERS-2 SAR in C-band. Thanks to their combination, multi-temporal studies in two bands with different polarizations were possible.

JERS-1 also carried two instruments with optical sensors: *Visible and Near Infrared Radiometer* (VNIR), *Short Wavelength Infrared Radiometer* (SWIR). These devices, operative in the visible, near and medium infrared, had difficulties in their functioning since the launch.

6.4.9.1 ADEOS-II

ADEOS (*Advanced Earth Observing Satellite* or Midori) is in some aspects similar to the Indian satellites IRS. It was launched on August 17, 1996 to an altitude of 800 km in sun-synchronous orbit and temporal resolution of 41 days. Its eight sensors have collected information on the heating of the globe, the thinning of the ozone layer and global climatic variations (Box 6.16).

The second of the series, Midori-II or ADEOS-II, launched on December 14, 2002, uses the same sensors as the first ADEOS for monitoring changes at global scale, and other more evolved sensors:

- AMSR, advanced microwaves radiometer, very similar to the device on board EOS-AM Aqua, in which JAXA contributes with AMSR-E;
- ILAS-II, spectrometer measuring the infrared radiation at the external edge of the atmosphere.

Box 6.16 Characteristics of the sensors on board the Japanese satellite ADEOS, Midori (JAXA)

Acronym	Sensor	Characteristics	Relevant applications
AVNIR	Advance Visible and Near Infrared Radiometer	Optical sensor 4 spectral bands: 0.42–0.50, 0.52–0.60, 0.61–0.69, 0.76–0.89 μm geometric resolution: 16 m 1 pancromatic band: 0.52–0.69 μm geometric resolution: 8 m Swath width: 80 km	Earth's surface and coastal zone observation in the visible and near infrared with good geometric resolution. Study for desertification, deforestation, coastal pollution, and estimation of environmental resources
OCTS	Ocean Colour and Temperature Scanner	Optical radiometer Geometric resolution: 700 m. Swath width: 1400 km Period: 3 days	Global data collection of the oceans physical characteristics concerning temperature and suspended materials
NSCAT	NASA Scatterometer	Radar scatterometer Geometric resolution: 25–50 km Monitoring of 90% of the Earth oceans every 2 days	Estimation of oceans wind velocity
TOMS	Total Ozone Mapping Spectrometer	Six bands of the near-ultraviolet	Earth albedo measurement, for O_3 and SO_2 estimation by means of absorption spectra registered by the sensor
POLDER	Polarization and Directionality of the Earth′s Reflectance	Electronic radiometer-polarimeter, developed by CNES	Study of polarization, direction and spectral characteristics of the suspended particles and aerosols in the atmosphere
IMG	Interferometric Monitor for Greenhouse Gases	Thermal infrared sensor	Monitoring of the terrestrial irradiative budget for green house studies
ILAS	Improved Limb Atmospheric Spectrometer		Spectral analysis of the stratospheric portion of the polar atmosphere to study ozone and chemical components over poles.
RIS	Retroreflector in Space	Intermittent laser at infrared rays	Quantitative atmospheric study

It was one of the four space platforms planned in simultaneous orbit with particular attention to the water cycle: EOS-AM Terra & Aqua, ADEOS-II and ENVISAT. But this satellite stopped its activity on October 25, 2003 due to faults in the electric system. The next collaboration includes the *National Polar-orbiting Operational Environmental Satellite System* (NPOESS), the United States' next generation of satellites to monitor the Earth's weather, atmosphere, oceans, land and near-space environment. JAXA is developing a new-generation microwave radiometer, the *Conical Scanning Microwave Imager/Sounder* (CMIS) to produce microwave imagery and other meteorological and oceanographic data.

6.4.9.2 ALOS

ALOS (*Advance Land Observing Satellite*) covers the fields of applications in mapping, precise regional land cover observation, disaster monitoring, and resource surveying. It represents the evolution of JERS-1 (Fuyo) and ADEOS (Midori) programmes on Earth Observation (Box 6.17).

Box 6.17 Optical and radar instruments on board the ALOS satellite

Acronym	Sensor	Characteristics	Relevant applications
PRISM	Panchromatic Remote sensing Instrument for Stereo Mapping	Panchromatic radiometer (0.52–0.77 μm) geometric resolution 2.5 m swath width 70 km at nadir and 35 km elsewhere	Three independent optical acquisition systems (nadir, forward, backward along the flight line) to generate stereoscopic images with high accuracy
AVNIR-2	Advance Visible and Near Infrared Radiometer type-2	Geometric resolution 10 m swath width 70 km	Multispectral sensor for land use/land cover mapping
PALSAR	Phase Array L-Band Synthetic Aperture Radar	All-weather conditions radar sensor in band-L (1.270 GHz)	Radar images with polarization HH, VV, HV and VH, incidence angle range 8°- 60°, geometric resolution from 7 to 44 m, radar antenna size: 8.9 m × 3.1 m.

ALOS (Daichi) carries three sensors:

- *Panchromatic Remote Sensing Instrument for Stereo Mapping* (PRISM), with three sets of optical systems to measure precise land elevation;

- *Advanced Visible and Near Infrared Radiometer* type 2 (AVNIR-2), for land cover analysis;
- *Phased Array type L-band Synthetic Aperture Radar* (PALSAR), enabling day-and-night and all-weather land observation.

6.4.9.3 TRMM

The *Tropical Rainfall Measuring Mission* (TRMM) is an EO satellite that diagnoses conditions by measuring precipitation in tropical and sub-tropical areas. TRMM is a joint project between Japan and the United States, launched on November 28, 1997 from the Tanegashima Space Center (TNSC). Japan provided the launcher and developed the Precipitation Radar (PR), while NASA developed the spacecraft, additional observation instruments, and the satellite operation systems.

6.4.10 CSA (Canada) Space Programme

6.4.10.1 Radarsat

The *Canadian Space Agency* (CSA), in French *Agence Spatiale Canadienne* (ASC), in the domain of Earth observation has concentrated its efforts in the development of synthetic aperture radar instruments.

The first Canadian satellite for EO, Radarsat-1, launched on November 4, 1995, uses SAR (*Synthetic Aperture Radar*) active microwave sensor in C-band and HH polarization for acquisitions with seven different depression angles from 20° to 49° with different ground resolutions, ranging from 10 to 100 m. The field of applications is wide, including cartographic restitution at scales from 1:50,000 to 1:5,000,000. The Radarsat-1 is characterized by a polar sun-synchronous orbit which enables the satellite to be independent of batteries. It orbits at 798 km altitude and takes 100.7 min to complete one orbit (14 orbits a day). The SAR antenna is 15 m $\times$ 1.5 m.

The Radarsat-2, launched on December 14, 2007, is characterized by several polarizations (HH, VV, HV and VH) and spatial resolutions from 3 to 100 m (Plate 6.6). The possibility to select the polarization offers different information:

- HH polarization: more suitable for studying soil moisture content;
- VV polarization: recognizes the land cover structure and the phenological conditions of the vegetation;
- HV and VH polarizations: not much sensitive to crop types, provide data about soil moisture and vegetation structure.

These experiences were possible thanks to some specific missions of the Space Shuttle described later in this chapter.

6.4.11 KARI (South Korea) Space Programme

The Korean medium-long term roadmap of the *Korean National Space Development Plan* (NSP) planned launches of KOMPSAT satellites of the next generations up to KOMPSAT-7.

The satellite KOMPSAT-1 (*Korean MultiPurpose Satellite*) was successfully launched on December 21, 1999; it was developed by the *Korea Aerospace Research Institute* (KARI), equipped with *optical–electronic camera* (EOC) and *Ocean Scanning Multispectral Imager* (OSMI). The satellite has sun-synchronous orbit (passing at 10:50 a.m.) at an altitude of 685 km; its weight is 500 kg.

Imagery acquired by the sensor EOC is suitable for producing cartographic maps at scale 1:25,000. The main mission of the first KOMPSAT is mapping Korea with a two-camera system. These sensors are characterized by an electronic array acquisition system providing panchromatic images (0.512–0.730 μm) with 6.6 m ground resolution and 17 km swath width. EOC can also record images with a ± 45° orientation, ensuring their overlap to provide a stereoscopic vision. As well as cartographic production, it permits the assessment of damages caused by earthquakes, avalanches, floods and to carry out a systematic monitoring to forecast any catastrophic phenomena.

OSMI sensor acquires colorimetric measures of the oceans with the purpose of managing and monitoring the ocean environment's resources. Each wavelength interval of the instrument is suitable for detecting a specific parameter in the marine environment (Table 6.13).

On July 28, 2006 the second *Korean MultiPurpose Satellite* (KOMPSAT-2) was launched by Eurockot from Plesetsk Cosmodrome, Russia. This represents an important step on the roadmap of the NSP, and an important addition to the already available high-resolution optical sensors. With specifications similar to those of IKONOS-2, it is also complementary to the remote sensing systems currently operated by Spot Image.

The metric satellite KOMPSAT-2 records data by the MSC (*Multispectral Camera* developed in collaboration with the *Israelian Space Agency*, in one panchromatic and four multispectral bands at resolution of 1 and 4 m respectively. The swath width is 15 km and the radiometric resolution 10 bit. The repeat cycle is 28 days which can be lowered to 1 – 3 days using body pointing of the satellite bus.

Table 6.13 Marine environment monitoring with the OSMI spectral bands on board KOMPSAT-1 (Korea)

Band centre (μm)	Characteristics
B1: 0.412	Suspended material
B2: 0.443	Chlorophyll concentration
B3: 0.490	Pigment concentration
B4: 0.555	Turbidity
B5: 0.765	Atmospheric effects calibration
B6: 0.865	Atmospheric effects calibration

6.4.12 The China–Brazil Cooperative (CBERS) Space Programme

The Brazilian space programme has significant capabilities in vehicles launching and satellite manufacturing. The reference body is the *Brazilian Space Agency* (*Agência Espacial Brasileira*, AEB).

The *China National Space Administration* (CNSA) has developed different types of man-made satellites, covering several fields in EO and telecommunication:

- the recoverable remote sensing satellites, DFH (Dongfanghong, or The East is Red) telecommunications and broadcasting satellites;
- FY (Fengyun, or Wind and Cloud) meteorological satellites;
- SJ (Shijian, or Practice) scientific research and technological experiment satellites;
- ZY (Ziyuan, or Resources) Earth resource satellites; and
- Beidou (Plough) navigation and positioning satellites.

CNSA has also developed the Long March rockets with several consecutive successful flights since 1996.

China and Brazil collaborate in a common EO space programme.

The *China–Brazil Earth Resources Satellite Program* (CBERS or ZY) was born in 1988, following the agreement of collaboration between the space agencies of the two countries for environmental studies.

The satellites CBERS-1 (October 1999–August 2003), CBERS-2, launched on October 21, 2003, and CBERS-2B, launched on September 19, 2007 from the base at Taiyuan (China), carry a recording data multi-sensor system provided with different instruments at different resolution (Table 6.14):

Table 6.14 Characteristics of the sensors on boards CBERS-1 and 2, Chinese–Brazilian satellites

Parameter	Unit	WFI	CCD Camera	IR-MSS
Spectral bands	μm	0.63–0.69 0.76–0.90	0.51–0.73 (pan) 0.45–0.52 0.52–0.59 0.63–0.69 0.77–0.89	0.50–1.10(pan) 1.55–1.75 2.08–2.35 2. 08–2.35
Field of view	degree, °	60	8.3	8.8
Geometric resolution	m	260	20	80 (Pan and Near IR) 160 (Thermal)
Temporal resolution	days	3–5	nadir: 26 3 days (± 32°)	26
Swath width	km	900	113	120

- *Wide Field Imager* (WFI) camera collects data of 260 m geometric resolution and 890 km wide swath in two spectral bands: 0.63–0.69 μm (red) and 0.77–0.89 μm (infrared);
- *High-Resolution Camera* (HRC) operates on a swath width of 120 km with possibility of ± 32° lateral vision for stereoscopic acquisition, 20 m spatial resolution, in five spectral bands: 0.51–0.73 μm (panchromatic); 0.45–0.52 μm (blue); 0.52–0.59 μm (green); 0.63–0.69 μm (red); 0.77–0.89 μm (near infrared). The lateral vision allows the system to reduce the temporal resolution from 26 days (nadir operation mode) to 3 days (off-nadir operation mode).
- *Infrared Multispectral Scanner* (IRMSS) operates in four bands 0.50–1.10 μm (panchromatic); 1.55–1.75 μm (infrared); 2.08–2.35 μm (infrared) and 10.40–12.50 μm (thermal infrared), with spatial resolution of 80 m (120 m in the thermal band) and synoptic field of view of 120 km.

CBERS-2B carries also the *High-Resolution Panchromatic Camera* (HRPC) with one single panchromatic band 0.50–0.80 μm which comprises part of the visible and of the near infrared portion of electromagnetic spectrum. The images recorded by this camera are 27 km width and have 2.7 m spatial resolution. A total of 130 days are required to obtain a full Earth coverage.

CBERS series is characterized by sun-synchronous orbits, at an altitude of 778 km, with 14 revolutions per day.

China has launched two new satellites, the Shi Jian-6 Group-03 (SJ-6E Shi Jian-6E and SJ-6F Shi Jian-6F), from its Taiyuan Satellite Launch Center on September 6, 2008. The Shi Jian-6 Group-03 satellites replaced the Shi Jian-6 Group-02 satellites launched on October 23, 2006.

The new satellites contribute to enhance experiments, monitoring the environment and forecasting natural disasters.

6.4.13 CONAE (Argentina) Space Programme

The Argentinian Space Agency, *Comision National de Actividades Espaciales* (CONAE), collaborates with several other Space Agencies in several fields.

In this context, three *Scientific Applications Satellites* (SAC) have been launched:

SAC-A: technological demonstration satellite;
SAC-B: devoted to scientific research;
SAC-C: the first Earth Observation Satellite of Argentina.

SAC-A was a small, ejectable, non-recoverable, low-cost satellite launched during the STS-88 Space Shuttle Endeavour mission, which included a *Differential Global Positioning System* (DGPS), a *Charge-Coupled Device* (CCD) camera and a magnetometer.

SAC-C is planned for multispectral imaging of terrestrial and coastal environments, studying the structure and dynamics of the Earth's atmosphere, ionosphere and geomagnetic field. With these objectives, CONAE has joined efforts with NASA in the creation of the *First International Earth Observation Constellation,* with Argentine SAC-C and the U.S. Landsat-7, EO-1 and EOS-AM Terra missions.

The programme SAOCOM, studied and planned by CONAE, includes the use of a Synthetic Aperture Radar SAR as a main acquisition tool. The first mission, SAOCOM-1, constituted by two satellites (1A and 1B), belongs to an Italian–Argentinian collaboration programme, SIASGE (*Sistema Italo-Argentino de Satélites protects Gestión de Emergencias*), which includes a connection with COSMO/SkyMed constellation.

The main radar instrument is in L-band, 23 cm wavelength, allowing a greater penetration than SkyMed and ERS/Envisat respectively in X- and C-bands in the leaf biomass and in soil moisture measurement and for water flow determination.

SAR L-Band geometric resolution can be programmed from 7 to 100 m and the swath width ranges from 50 to 400 km in function of the selected operational mode. The imagery recording capacity on board is up to 100 Gbit, and it can transfer data to the receiving station at a speed of 150 Mbit/s for each of the two available channels.

6.4.14 International Space Station (ISS)

In the context of the description of the systems for Earth observation, it is opportune to give particular attention to the most complex international mission in history, the *International Space Station* (ISS). This project was first announced in 1993 and called *Space Station Alpha*. It combines the scientific collaboration of 16 countries (Canada, Japan, Russia, Brazil, USA, Italy, Belgium, Holland, Denmark, Norway, France, Spain, Sweden, Switzerland, United Kingdom and Germany) each of them in charge of one or more modules of the mission. The satellite platform is constituted by many components, assembled in space starting from November 1998.

The station is serviced primarily by Russian *Soyuz* and *Progress* spacecraft and by U.S. *Space Shuttle* orbiters with crew. The *European Space Agency* (ESA) launched on March 9, 2008, *Ariane-5* with the first *Jules Verne Automated Transfer Vehicle* (ATV) towards the ISS carrying over 8000 kg of cargo. Other vehicles are the Japanese (JAXA) *H-II Transfer Vehicle* (HTV) and the American (NASA) *Orion*.

For the completion of Space Station, 45 missions are planned.

The station orbits at approximately 350 km of altitude with inclination of 51.6° and speed of 27,700 km/h, completing 15.77 orbits per day. At the end of the assembly operations, it will be 108 m wide and 88 m long weighing 450 tons, and it will be able to lodge astronaut crews performing scientific research in turns of 90 days (see Plate 6.4).

The ISS has been planned to carry out innovative research and diversified studies in the:

- *medical field*: thanks to the absence of gravity, in space it is possible to grow proteic crystals which will help in investigations about enzyme structures, viruses and, maybe to create new pharmaceuticals. Also studies about in vitro cell growth are carried out and analyses about the biological feedback of organisms in the absence of gravity will be investigated more deeply;
- *physics domain*: regarding material combustion in conditions of zero gravity. It also deals with analysing space particles and vacuum characteristics;
- *environmental sciences*: the Earth is observed from the station with the purpose to collect (spatial and temporal) large-scale data about terrestrial environment variations. Knowledge about forests, oceans and mountains can be increased and the effects of catastrophic events on many environmental aspects can be analysed. Attention is also directed to urban and marine pollution, deforestation and desertification.

Therefore, the range of research includes a wide spectrum of disciplines, such as biology, bio-medical, biotechnology, physics, fluid physics, material science, quantum physics, astronomy, cosmology, meteorology and Earth Observation.

6.4.15 Radar Missions on the Space Shuttle

The vehicle Space Shuttle, called *Space Transportation System* (STS), is the spacecraft used by the United States for its frequent human spaceflight missions into orbit around the Earth.

The first space vehicle was the shuttle *Columbia* which realized the first return expedition in April 1981 and was tragically destroyed when going back into the atmosphere in January 2003. NASA's space programme also uses the shuttles *Challenger*, *Discovery*, *Atlantis* and *Endeavour*. The next generation of Space vehicle is *Orion*, designed to take humans to the Moon.

The shuttle carries out experiments of different types and is actively involved in the realization of the *International Space Station*. Its missions are crew-equipped, at 250–350 km altitude orbits.

The first instrument used since 1981 is the high-precision calibration large format optical photo camera, used for photographic–cartographic acquisitions. Known as the *Large Format Camera* (LFC), this photo camera can record 225 km $\times$ 459 km (h: 300 km) areas of the terrestrial surface on metric film having an image format of 23 cm $\times$ 46 cm, and 10 m photographic resolution on black-and-white films (see Plate 6.5).

Many experiments to demonstrate the Shuttle's potentialities for research and studies about terrestrial resources by using radar tools were carried out:

- in 1981 the *Shuttle Imaging Radar* (SIR-A) experiment developed by the Seasat satellite L-band radar, with 40 m geometric resolution, was carried out to assess the orbital radar potential for geological maps production;

- in 1984 with SIR-B, differing from the previous mission for a better ground resolution (25 m) and for the 47° incidence angle, not fixed any more, like in SIR-A, but variable in the range 15°– 60°;
- in 1994 the SIR-C/X-SAR, with three bands L, C and X and VV, HH, HV, VH polarizations (Plate 6.6);
- in 1997, mission similar to the previous one for the use of multi-frequency and multi-polarization sensors;
- from 12 to 23 February 2000, the *Shuttle Radar Topography Mission* (STRM), acquiring interferometric images. Collected data were used to produce a map of the Earth's surface in the latitude interval +60° and –56°. The principle used was radar interferometry by the use of two antennas: the first one fixed on the Shuttle and the second one set at the end of a pool to be opened when the planned orbital altitude is reached. The *simultaneous* use of two antennae gets acquisitions of the same scene from two different angles of view producing coherent interferometric images in radar band, with the possibility to show the third dimension, i.e. the height. The global SRTM Digital Surface Model (DSM) data set has 90 m postings, but it has several holes, where the data are not-existents for several reasons.

Studies of these images confirm that the multiple combination of radar data in different polarizations and wave frequencies provide an interesting new approach to the knowledge of the terrestrial resources and morphology (Table 6.15).

DSMs are important for creating a number of products from other satellite sensors. The build of a better Global DSM is scheduled by USGS and Earth Remote Sensing data Acquisition Centre in Japan cooperating in a fused ASTER/SRTM data set to give 30 m postings, using the Indian CartSat optical satellite to fill in voids. Such a DSM has already been produced in India.

6.4.16 Commercial Satellites

Traditionally remote sensing systems have been used for global, regional and national programmes as support to government activities. Now new geospatial markets are emerging. The authorization for the property and the activity, in practice with no restrictions, of satellites for the Earth observation has stimulated competitiveness among the largest aerospace worldwide companies to create new geospatial products and services required by the market.

During the apex of the Cold War, the governments had control of the remote sensing satellites restricting the geometric resolution at not lower than 10 m for Earth observation civilian satellites. This constraint has been removed opening the resolution below 1 m for civilian and commercial applications.

The systems like Landsat (USA), SPOT (France), IRS (India), Spin and DK (Russia), ERS (Europe), Radarsat (Canada) and others are largely financed by the gov-

Table 6.15 Characteristics of the main radar missions (see Box 6.5)

Satellite	Unit	Seasat-A	Sir-A	ERS-1 and 2	ENVISAT	Space Shuttle	PALSAR	RADARSAT-2
Sensor	–	SAR	SAR	SAR	ASAR (image)	SIR-C/X-SAR	SAR	SAR
Country	–	USA	USA	ESA	UE (ESA)	USA, D, I	Japan	Canada
Operativeness	–	1978	1981	1991(95) op	2001 op	1994	2004	2000 op
Band	–	L	L	C	C	X, C, L	L	C
Wavelength	cm	23	23	5.66	5.33	3- 5.6–23	23	5.17
Frequency	GHz	1.270	1.270	5.300	5.331		1.270	5.405
Polarization	–	HH	HH	VV	VV HH	VV HH HV VH	VV HH HV VH	VV HH HV VH
Slant angle	°	20–70	47	23	23	15–60	8–60	20–49
Orbit altitude	km	790	–	790	600	250	568	790
Period	days	1 hour 40’	–	35	35	–	44	14
Ground resolution	m	25	40	25	30	25	7–44	3–100
Swath width	km	100	8	100	100	8 (with 23°)	75	20–500

ernments (more than 90%) as their main activities are of public utility. Commercial satellites are strongly financed by the industry.

The expansion of the market is based on the capacity of the sensors carried by the satellites to provide in-time data that can reply to the requests of the end user (geospatial information, geo-marketing, real estates, insurances, etc.), as well as satisfy the needs of the traditional cartography and of the scientific community. About this topic, there are many issues like:

- government restrictions imposition;
- international variations in the rights law system;
- prices equity;
- competition with government programmes;
- necessity of operational self-control;
- data and information quality;
- teaching and training.

The market for geospatial information guarantees a significantly better spatial and temporal accuracy than governments' accuracy, but spatial and spectral quality of the images acquired by many and different commercial systems still have to be verified.

In particular the uncertainty about the image quality is high: traditionally the governments have supported the development of standards and specifications for images' metadata as they themselves have been the main consumers, for example for cartography, meteorology, environmental protection, civil protection and defence. The development of many systems in competition with each other leaves the issue open, and in the free market a larger offer is expected.

Attention is given also to the geometric accuracy (Table 6.16). Commercial systems have geometric resolution from 0.5 m up in panchromatic and from 2.0 m up in multispectral. The planning of missions, orbits and technical solutions, like off-nadir view up to 45° along the flight line and lateral, generates a higher frequency of acquisition reducing the revisiting period to 1–5 days and in some cases (EROS, COSMO/SkyMed) also more times in the same day.

In the multispectral, the tendency is to increase the number of bands and to improve the spectral resolution. The availability of a higher number of bands provides, as a consequence, a larger quantity of available information which has to be properly selected in function of the studies and applications to be performed. Not all the acquired and available bands need to be used, as some are correlated among them, i.e. provide very similar information, or as the application is completed by using specific bands (i.e. some bands in the visible and near infrared for studying crop differentiation and phenology).

Since 1997 several satellites with high resolution have been launched and many other systems with high, medium and low resolution are in preparation and scheduled.

Operative or operative in short time missions have the following main characteristics:

Table 6.16 Operative characteristics of commercial satellite

Satellite	Geometric resolution (m)	Geometric accuracy (m)	Scene frame (km × km)	Stereoscopy
SPIN-2 TK-350	10	–	200 × 300	Yes
SPIN-2 KVR-1000	2	20	40 × 160	No
QuickBird	0.61	2	36 × 36	Yes
OrbView-3	1	2	8 × 8	Yes
GeoEye-1	0.41	2	15.2 × 15.2	Yes
IKONOS	1	2	11 × 11	Yes
EROS-A1	1	2	8 × 8	Yes
DK-1	1	5	28.3 × 28.3	Yes
DK-2	2	5–10	28.3 × 28.3	Yes
COSMO/SkyMed radar	3	3–6	from 24 × 24	Yes
COSMO/SkyMed Pleiades	0.7	2	from 15 × 15	Yes

- immediate operativeness with no test phases;
- flexibility in planning the acquisitions and in the pointing defined by the user;
- high resolution and geometric accuracy;
- regional acquisition of reduced size scenes due to the high geometric resolution;
- high temporal resolution;
- possibility of stereoscopic vision along and on the side the same path (with short delay of acquisition between the images);
- acquisition and delivery of pre-processed images even within 24 h;
- tendency to plan satellites' constellations.

The series and/or constellation of high-resolution satellites IKONOS, EarlyBird, OrbView and EROS are briefly described here highlighting the main characteristics of each one in respect of its limits (Table 6.17). They have two sensors on board, panchromatic (1 band) and multispectral (4 bands). They can perform stereoscopic acquisition both on the flight line and obliquely with angles ranging from ± 30° (IKONOS, EarlyBird, OrbView) to ± 45° (EROS).

They complete 14 orbits per day with 98 min period. All of them pass at fixed time 10:30 a.m., equator time, unless otherwise specified. The dynamic range is usually 11 bit (2048 levels). The comparison between the parameters is reported in Table 6.19.

6.4.16.1 IKONOS

From the Greek word for image, IKONOS is the first real remote sensing commercial satellite operative in orbit since September 24, 1999 with 1 m geometric resolution (Fig. 6.27).

IKONOS has 12 m both horizontal and vertical nominal accuracy without GCPs and 2 m horizontal and 3 m vertical accuracy with GCPs. The instruments have stereoscopic acquisition capacity along the flight line. This technical solution acquires an image with a different field of view with a small delay Δt compared to the previous one.

Table 6.17 Characteristics of the high-resolution satellite-sensor systems IRS-1D, IKONOS, QuickBird, OrbView, EROS

Parameter	Unit	IRS-1D	IKONOS	QuickBird	GeoEye-1	EROS-C
Temporal resolution	days		1–3	3–4	2–8	
Equator altitude	km		681	450	684	500
Dynamic range	bit		11	11	10	10
Scene size	km^2		13×13 11×100 11×1000	16.5×16.5	15.2×15.2	7×7
Sensor			PAN	PAN	PAN	MSPAN
Spectral resolution						
Panchromatic	μm		0.45–0.90	0.45–0.90	0.45–0.90	0.50–0.90
Blue	μm		0.45–0.52	0.45–0.52	0.45–0.52	0.45–0.52
Green	μm		0.52–0.61	0.52–0.60	0.52–0.60	0.52–0.60
Red	μm		0.64–0.72	0.63–0.69	0.63–0.69	0.63–0.69
Near IR	μm		0.77–0.88	0.76–0.90	0.76–0.90	0.76–0.90
Geometric resolution	m		1 4	0.61 2.44	0.41 1.6	0.70 2.40
Pixel / line			13,000	27,424		20,000
Pixel per km^2			1.10^6	$1.64.10^6$		

Fig. 6.27 Window 400×200 pixels, IKONOS panchromatic band, Livorno area, Tuscany, Italy

Potential applications of high-resolution data are suitable for different types of users. The accuracy and the easy to read images give ideal support for mapping and analysis activities. The data can be rendered by pan-sharpening technique combining 1 m spatial resolution of panchromatic data with the information of 4 m multispectral data.

The following high-resolution products are provided:

- *georeferenced*: geometric corrections and rectification on an ellipsoid specified by the user and projected on a map. This product has ± 50 m horizontal accuracy;
- *orthorectified*: precision products suitable for territorial and urban mapping and GIS applications requiring positioning accuracy. They are realized through ground surveys and digital elevation models. The best products have ± 4 m horizontal accuracy compatible with 1:5000 scale maps.
- *3D vision*: visualizes areas in three dimensions. This interactive application provides an intuitive control with virtual rendering over the scene visualizing transport infrastructures, mountains, valleys, watersheds and other natural and entropic features.

6.4.16.2 QuickBird

This high-resolution project encountered two failures: the launch of EarlyBird satellite, 3 m resolution panchromatic and 15 m multispectral (MS), first, and later the QuickBird, 0.82 m panchromatic and 3.28 m multispectral.

In December 2000 Earth Watch, the society that manages the programme, obtained the licence from NOAA to operate with resolution until 0.5 m.

Reducing the altitude of the sun-synchronous orbits to 450 km, QuickBird-2, successfully launched on October 18, 2001, acquires with ground resolutions of 0.61 m in the panchromatic and 2.44 m in the multispectral. In 25° oblique acquisition, the geometric resolution is reduced respectively to 0.72 m Pan and 2.8 m MS, however giving the opportunity of stereoscopic acquisition (Plate 6.3).

The period for passing over the same area at 40° latitude north is 3.5 days, using the lateral pointing system.

The swath width is 16.5 km at the nadir. The acquisitions can be both in 10 image stripes (165 km) and in 2×2 scene areas recorded during one only passing. The acquisition time of one scene (16.5 km × 16.5 km) is 4 s.

6.4.16.3 OrbView

Orbimage (now GeoEye) is a consortium of companies specialized in satellite launch and manufacture, controlling the platforms OrbView-1, an atmospheric satellite launched in April 1995, and OrbView-2, Ocean and Land Satellite, low (1.1 km) geometric resolution satellite launched in 1997 (Plate 6.7); the third and fourth

Table 6.18 Characteristics of the OrbView 1, 2, 3 and 4 satellites

		OrbView series			
Parameter	Unit	OrbView-1	OrbView-2	Orbview-3	Orbview-4
Geometric resolution	m	10 km	1000	1 Pan 4 multispectral (MS)	1 Pan 4 multispectral 8 – 20 (hyperspectral)
Spectral resolution	nm	1 band: 777	8 bands: 402–885	4 bands (MS):450–900 1 panchromatic band	1 band (Pan) 4 bands (MS) 200 bands (hyperspectral)
Swath width	km	1300	2800	8	8 km (PAN and MS) 5 km (hyperspectral)
Orbit altitude	km	740	705	470	470
Mission period	years	5	7.5	5	5

series, OrbView-3 (launched in June 26, 2003) and 4, includes OrbView series' high-resolution satellites (Table 6.18).

GeoEye′s OrbView-2 satellite with the SeaWiFS sensor provides climatology of global chlorophyll and other Environmental Data Records (EDRs) important for climate change and global warming studies since 1997.

OrbView-4 was a high-resolution commercial Earth imaging satellite. OrbView-4's imaging instrument provided 1 m panchromatic imagery and 4 m multispectral imagery with a swath width of 8 km as well as 200 channel hyperspectral imagery with a swath width of 5 km. The satellite revisited each location on Earth in less than 3 days with an ability to turn from side to side up to 45° from a polar orbital path.

OrbView-4 was lost in a launch failure on September 21, 2001, when the carrier rocket suffered a loss of control and the orbit was not achieved.

OrbView-5 (GeoEye-1), launched September 6, 2008, orbits 684 km above the Earth in a sun-synchronous path while collecting imagery at 0.41 m resolution in the panchromatic (black-and-white) mode, and 1.64 m resolution in the multispectral (colour) mode. The system has full imaging and stereoscopic capacities in any direction.

6.4.16.4 EROS

The constellation of low orbit satellites, *Earth Remote Observation System* (EROS), derived from Israelian military technologies, represents another opportunity for the use of high geometric resolution images, especially for the possibility of daily and different time acquisition of the same area and the acquisition of stereoscopic pairs along swathes up to hundreds of kilometres long.

The project is managed by *West Indian Space* (WIS), according to a licence released by the Israelian government. The programme, part of ImageSat vision,

includes the launch of a constellation of high geometric resolution satellites divided into three series:

- EROS-A: panchromatic device with 1.8 m geometric resolution;
- EROS-B: panchromatic instrument with better resolution (0.7 m);
- EROS-C: high geometric resolution panchromatic and multispectral acquisitions.

The constellation will guarantee daily acquisition frequencies for each point of the Earth thanks to the possibility of re-orienting the sensors of $\pm$ 45° and will also be able to acquire in different solar illumination conditions (equator crossing: 10:00, 10:30, 11:00, 13:00, 13:30, 14:00).

Some plan parameters in A and B series of EROS constellation are different. In EROS-A the geometric resolution of the panchromatic instrument is 1.9 m. For nadir-looking geometry, the EROS-A satellite delivers standard panchromatic images in 14 km swath with 1.8 m resolution, while in the hypersampled mode it provides customer-specified image acquisitions in reduced 9.5 km swath and with an oversampled 1.2 m resolution.

EROS-B have superior capabilities compared to EROS-A, including a larger synchronous camera of CCD/TDI type (*Charge-Coupled Device/Time Delay Integration*), like QuickBird, IKONOS and OrbView-3, with standard panchromatic resolution of 0.70 m at an altitude of about 500 km, larger on-board recorder, improved pointing accuracy and faster data communication link. The synchronous camera restricts the geometric deformations and the ground resolution changes.

The EROS-C satellite will be equipped with a camera with CCD/TDI sensors, producing both panchromatic imagery of 0.70 m resolution and multispectral imagery capabilities of 2.8 m resolution, with a swath of 11 km at nadir, offering higher quality resolution and higher data link rate.

The characteristics of supply of the images are particularly interesting; the scenes' shot, as well as in IKONOS, is different than the conventional one used in all the other images. Imagery, in fact, can also be required for long swathes (even hundreds of kilometres) as wide as the acquisition or for mosaics of many scenes (9, 18, 36...), as well as for stereoscopic swathes, with different modes among the satellites of the constellation.

EROS-A1 was launched on December 5, 2000 and EROS-B1 on April 25, 2006.

6.4.16.5 Rapideye

Rapideye fleet of five identical satellites has been launched August 29, 2008 on a DNEPR-1 launch vehicle from the Baikonur spaceport in Kazakhstan. This constellation is in a shared orbit at 630 km and circles the globe 15 times daily. Rapideye is able to deliver high-quality data for commercial use in land planning, agriculture, forestry and cartography. The satellites image any area in the world at latitude between below 75° north and south within 1 day on demand.

The multispectral pushbroom imager on board each spacecraft images the Earth in five spectral bands, scanning a 77 km swath at 6.5 m resolution, each strip with a maximum length of 1500 km.

The German Aerospace Center (DLR) has supported Rapideye in the development of the constellation.

6.4.16.6 TerraSAR-X

The TerraSAR-X space system was proposed by the *German Aerospace Centre* (DLR) for implementation as an element of ESAś *Earth Watch* programme, and it includes active radar in band X (3 cm). It was launched June 15, 2007 by a Dnepr rocket from Baikonur, Russia.

The radar instrument will be able to acquire in X-band frequency and a maximum of 1 m geometric resolution.

The X-Band SAR delivers radar imagery in the following modes:

- Image Mode Spotlight, 1 m resolution;
- StripMap, 3 m resolution;
- ScanSAR, 16 m resolution.

The acquisitions, from the orbit altitude of 515 km, can be single and multi-angle varying the incident angle in the range from 23° to 60°.

The principal scientific objectives are

- land cover mapping and change detection for GMES services;
- forest certification and cover monitoring;
- precision farming and crop-type mapping;
- topographic mapping;
- regional and urban planning, environmental protection support.

With the endorsement of the future TanDEM-X mission, it will be possible to completely measure the Earth's land surface, which is 150 million km^2, within a period of only 2.5 years. For a 12 m grid, height information can be determined with an accuracy of < 2 m.

6.4.17 Other Missions

Other countries have also started Earth Observation space activities in cooperation with private and commercial initiatives. Among them, the *Disaster Monitoring Constellation* (DMC) consists of five remote sensing satellites operated by the Algerian, Nigerian, Turkish, British and Chinese Space Agencies. The DMC provides imagery in support of deforestation, disaster relief and agricultural monitoring.

The individual DMC satellites are

- AlSAT-1 (Algeria) launched November 2002.
- BilSAT (Turkey), which completed its mission in August 2006 due to failed battery cells; with three sensors: 12 m resolution panchromatic, 24 m resolution multispectral, 9 bands super-spectral;
- NigeriaSAT-1 (Nigeria)
- UK-DMC (United Kingdom), with the simultaneous launch, on September 27, 2003. UK-DMC, BilSAT and NigeriaSAT-1 micro-satellites uses three spectral bands similar to Landsat 7 bands 2, 3 and 4 and 32 m geometric resolution with an exceptionally wide swath width of over 640 km; UK-DMC-2 is expected with 22 m resolution.
- Beijing-1 (China) launched October 2005, carries two payloads that provide high-resolution (4 m) panchromatic images along with medium resolution (32 m) multispectral images with an ultra-wide 600-km imaging swath.

6.5 Airborne Systems

The advantage of aircraft acquisitions derives from the large flexibility, adaptable to every specific need, especially temporal. However airborne images have more distortions than satellite images mainly due to the instability of the acquisition platform. To have an adequate swath width, the sensors must have a wide IFOV, which generates dilatation of pixels in off-nadir acquisition. The geometric resolution depends on the IFOV and on the flight height (H). Their product gives the pixel size at nadir. For example, with IFOV: 2.5 mrad and H: 1 km the pixel side is 2.5 m. If the altitude increases to 10 km, the pixel side is 25 m. Lower flight heights thus allow smaller pixels on level ground.

Pixel geometric distortions due to the relative altimetry are more visible from aircraft than from satellites: i.e. the Alps height (~4000 m) from Landsat (~7500 km) is only ~0.5% of the orbital altitude.

Airborne acquisitions can be made by three main methods:

- traditional photographic and photogrammetric instruments for the production of stereoscopic pairs;
- one or more video cameras in the visible and/or in the infrared;
- digital multispectral scanners acquiring simultaneously the same scene in more spectral ranges in visible and near, medium and thermal infrared (Table 6.19).

The traditional photographic instruments and the video cameras can be fixed on an aircraft for acquisitions, observations and analyses in the interval of visible and/or photographic near infrared (up to 0.9 μm) and medium infrared (up to 2.0 μm for the video cameras).

The scales of representation in photogrammetric acquisitions are generally included between 1:2000 and 1:25,000 with special applications at smaller and larger scales. The synoptic view is a function of the flight height, and the geometric resolution and flexibility of acquisition are better than from satellites.

The metric film *KODAK 2444 III Aercolour Negative* of 23 cm is the photographic support used. It has the best characteristics of quality and a potential resolution at the maximum contrast not lower than 100 lines/mm; the support of the sensitive emulsion offers the best characteristics of unwarpage and sensitivity and grain's thinness which are optimal qualities for the photogrammetric flights. The variations of the distances between fiducial marks, in comparison with the calibration nominal marks, are less than 0.01 nm.

6.5.1 Aerophotogrammetric Digital Cameras

All the domains dealing with the acquisition of geospatial information are dependent on digital technology except the large format traditional photogrammetric domain, where photographic film is still the source of data.

However digital aerophotogrammetric cameras are evolving and already have a scientific and practical importance although they do not yet have all the evidence needed to replace the photographic film in its definite advantages.

The knowledge regarding films has reached the maximum, while digital techniques are in continuous evolution, and the possible technological and methodological solutions are progressing in their development.

Linear or *matrix* CCD acquisition techniques have now entered the domain of photogrammetry. Punctual scanning cameras are excluded due to the difficulties linked to the reconstruction of rigorous acquisition geometry.

Table 6.19 Radiometric characteristics of some multispectral, super-spectral and hyperspectral airborne sensors

Sensor	Spectral range (μm)		Resolution (μm)	Number of bands
Airborne TM	Visible-NIR	0.42–0.90	0.03–0.14	7
	Medium IR	0.91–2.35	0.14–0.30	3
	Thermal IR	8.5–13.0	4.5	1
CZCS	Visible-NIR	0.42–1.05	0.03–0.14	9
	Medium IR	1.55–2.35	0.20–0.30	2
	Thermal IR	8.5–13.0	4.5	1
Thermo camera	Thermal IR	8.5–13.0	0.1°C	1
MIVIS	Visible-NIR	0.433–0.833	0.02	20
	Medium IR	1.150–1.550	0.05	8
	Medium IR	2.000–2.500	0.008	64
	Thermal IR	8.200–12.70	0.4	10

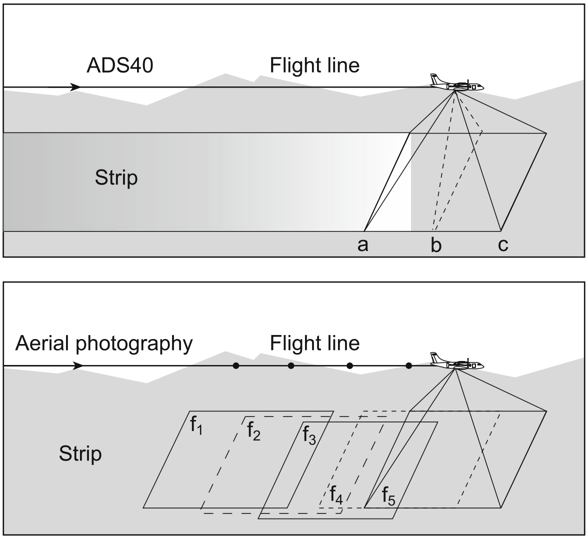


Fig. 6.28 Digital acquisition of the tri-linear ADS40 aerophotogrammetric sensor: (**a**) backward, (**b**) vertical (nadir), (**c**) forward, compared with photographic acquisition where 60% of the object within a scene is recorded three times. f_n: photogram number.

For linear or matrix modes, there are many contrasting points, reciprocal advantages and disadvantages. Each system has its characteristics for which it is suitable for some acquisitions and its applications:

- *linear*: medium-scale panchromatic acquisitions, colours and infrared false colour.
- *matrix*: digital aerophotogrammetric acquisitions for large-scale recording of urban scenes.

They are both helping to make the traditional photogrammetric geometric approach closer to remote sensing.

Among the photogrammetric cameras produced, there are some that have already been used for experimental and practical issues.

6.5.1.1 Airborne Digital Sensor (ADS40)

The digital photogrammetric camera ADS40, based on the principle developed by Otto Hofmann in the early 1970s, solves the problems of stereoscopic acquisition and data transfer velocity using three series of sensors set on linear arrays (Table 6.20). They acquire in the same moment, producing then stereoscopic terns along the swath, and using adequate systems which guarantee a speed of data transfer of about 50 Mbytes/s.

Linear sensors acquire the whole image along the flight line, or swath, using the platform motion in three different positions: vertical, or nadir, skewed forward and backward with respect to the aircraft instantaneous position (Figs. 6.28 and 6.29).

The sensors in the panchromatic (Pan) is made by 24,000 elements, or cells, arranged on two rows of 12,000 elements each, half a cell shifted one in respect to the other.

The multispectral sensor (MS) arrangement includes another four 12,000 cell linear elements, each one provided with filters for acquisition in the visible and near infrared.

The resolution is about 7 μm (the pixel's size on the ground depends on the flight height) with a positioning accuracy of ±1 μm. The system is equipped with a kinematics GPS integrated in the instrument and with an inertial platform (INS) which guarantees the reconstruction of each scanning line's geometry (Fig. 6.30).

In theory the triple acquisition of the same object from three different angles of view can generate a triangulation of point to produce a Digital Terrain Model (DTM).

6.5.1.2 Modular Digital Camera (DMC)

DMC (*Digital Modular Camera*), officially presented at ISPRS Congress in Amsterdam in July 2000, is a digital camera acquiring in panchromatic, one band, and multispectral, in four spectral bands with different geometric resolutions. Table 6.21 shows the main technical characteristics.

Unlike ADS40, the DMC operative principle is based on the reconstruction of a central perspective, like in the traditional photogrammetric camera. Three different rectangular matrices of sensors are used, synchronized among themselves, simultaneously acquiring three portions of the photogram reconstructed in its whole by a complex mosaic operation (Plate 6.8).

The advantage of this camera is represented by image acquisition similar to the traditional analogical photogrammetry, ensuring the direct use of digital photogram-

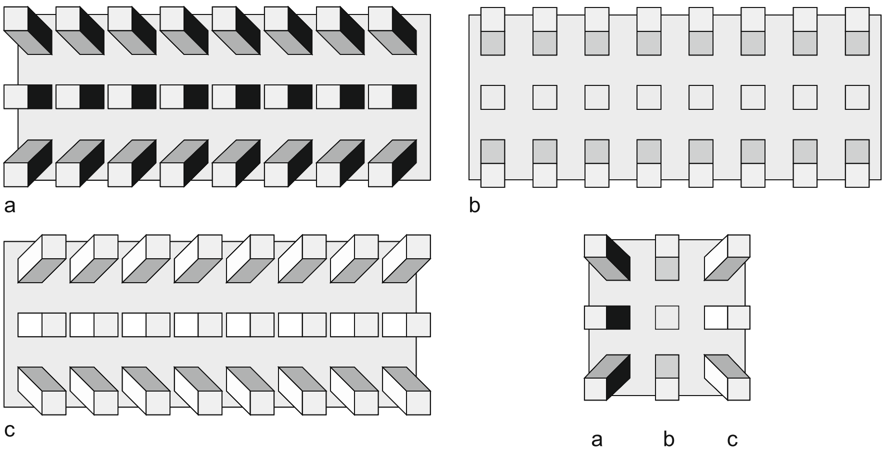


Fig. 6.29 Geometric distortion of objects in a scene acquired along the flight line: (**a**) backward, (**b**) vertical, (**c**) forward

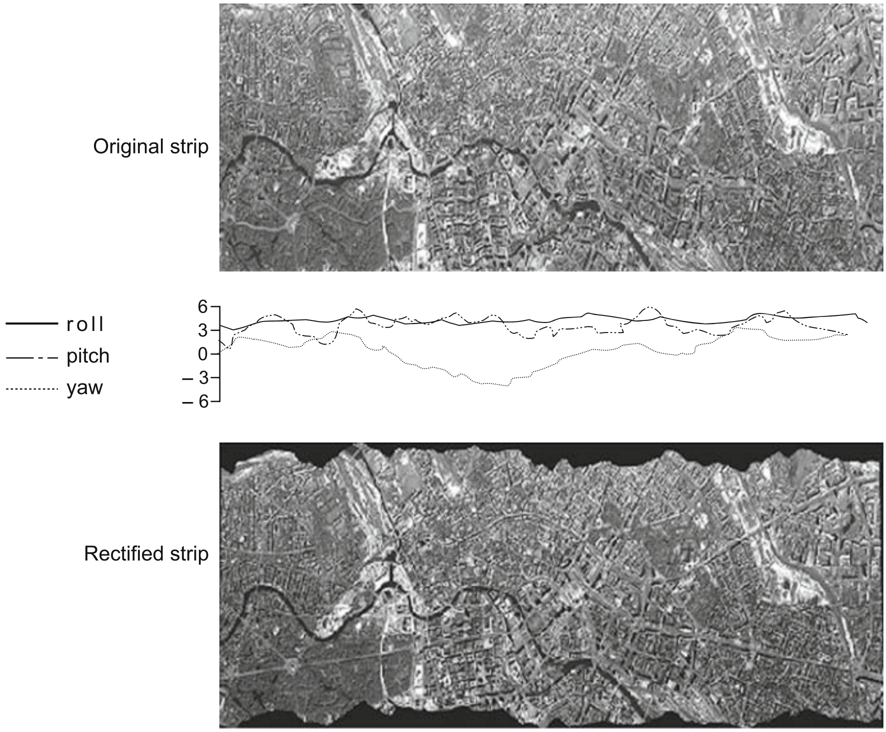


Fig. 6.30 Swath simulation of the sensor ADS40 with post-processing that uses the inertial platform data correction to determine the three distortion pitch, yaw and roll that generates the angles φ, ω and k

metric projection software. However it is important to know the modes of scene reconstruction, by geometric and radiometric interpolation, in order to better evaluate the quality of the acquired images.

Tests of acquisition, in optimal conditions, show the possibility to acquire pixels of 7 cm geometric resolution using the following flight parameters:

- sensor cell size: 12 μm
- lens focal length 50 mm
- height of acquisition 300 m
- advancing velocity 70 m/s

At height of acquisition over 1500 m, it is possible to obtain with digital cameras geometric resolution of the order of some decimetres.

6.5.1.3 High-Resolution Stereo Camera-Airborne (HRSC-A)

The *High-Resolution Stereo Camera-Airborne* (HRSC-A) digital camera, at the beginning developed for the missions of Mars exploration, is of linear type (pushbroom).

Table 6.20 Technical characteristics of the photogrammetric digital camera ADS40

Parameter	Unit	Characteristics
Sensor elements	n°	Panchromatic 24,000 multispectral 12,000
Spectral range	μm	Panchromatic 0.46–0.68 multispectral B: 0.43–0.49 V: 0.54–0.59 R: 0.61–0.66 VIR: 0.84–0.88
Geometric resolution	μm	7
Dynamic range	bit	12 (raw data)
Radiometric interpolation method		Linear 8 bit
Field of view (FOV)	degree, °	64
Compression factor		2–20
Data rating	msec	> 1.2
Spectral resolution		Panchromatic, visible and near infrared

Table 6.21 Characteristics of the possible configurations of the digital camera DMC (Hinz, 2000)

Configuration parameter	Unit	14 K pan/MS	7 K colour	7 K Pan	3 K MS
Combined digital cameras	n°	4	2	1	1
Field of view (FOV)		74° × 44°	39° × 42°	39° × 22°	72° × 50°
Geometric resolution Pan n° of pixel into the matrix	n°	13,500×8000	7000×7500	7000×4000	7000×4000
Geometric resolution MS n° of pixel per band	n°	3000×2000×4	3000×2000×3		3000×2000×4
Lens aperture and focal length		variable	variable	variable	variable
Memory:					
- Gigabyte	Gb	840	560	280	280
- images	n°	3000	3700	5000	5800
Acquisition time per scene	s	2	2	2	2
Dynamic range	bit	12	12	12	12
Weight	Kg	80	65	50	50

The multispectral and multi-stereo nine-line array instrument obtains orthoimages and DTM with accuracy of the order of 10–20 cm.

The linear sensor is made by 5184 CCD detectors, of which nine are parallels arranged on the focal plane of the instrument. The first five are panchromatic detectors set to have five different angles of view allowing the instrument to have a multiple stereoscopic vision. The other four linear arrays are covered by selective filters acquiring three bands in the visible and one in the infrared.

In this camera, as well as in ADS40, the presence of a Global Positioning System (GPS) and an Inertial Navigation System (INS) is necessary in order to continuously control the movements of the aerial platform and correct the acquired data.

The volume of collected data is huge. An example: an area of 2 km $\times$ 2 km can generate 32 Gbytes of input data and 3.2 Gbytes for DTM and orthoimage. As a comparison with the volume of data collected by a satellite sensor, a 15 km $\times$ 15 km area acquired by this digital camera generates 100 times more data than in a 60 km $\times$ 60 km SPOT image.

Each acquisition by the HRSC-A stereo-camera provides the following data and products:

- high-resolution panchromatic data with five different angles of view of the same scene;
- high-resolution multispectral data in the visible (3) and near infrared (1);
- Digital Terrain Model.

At 6000 m altitude, the ground pixel size is 24 cm and the accuracy $\pm$ 20 cm in latitude (x) and longitude (y) and 30 cm in height (z). At 70 m/s speed the data acquisition level is 7 Mbytes/s (5184 pixels/scanning line $\times$ 450 scanning lines $\times$ 3 bands). The storing capacity is 50 Gbytes corresponding to 2 h of acquisition.

6.5.1.4 Linear or Matrix

Both the solutions in linear acquisition and matrix modes have advantages and disadvantages and both are not suitable to completely replace photographic film. As the technologies are completely different between them, it is not even possible to join the advantages to obtain a unique optimal instrument (Box 6.18).

The three-linear digital camera is more appropriate for applications typical of remote sensing, as it absorbed its basic concepts, whereas the matrix camera is more suitable for better precision topographic restitutions (Box 6.19).

The use of positioning (GPS) and inertial (INS) systems are necessary for the three-linear system, useful but not essential for the matrix system.

The data acquired by the new sensors can be used for many applications:

- restitution producing numerical vector cartography;
- restitution of photo-planes and orthophoto-planes, non-vector product;
- extraction of thematic information analysing the texture characteristics of images;
- extraction of thematic information from multi- and super-spectral images;
- 3D representation at large scale of urban areas.

Box 6.18 Strength and weakness of the linear and matrix photogrammetric acquisition systems and relevance in defined conditions

Acquisition system	
Linear	**Matrix**
Strength	
Process fully digital	
Higher radiometric quality of the digital sensors compared with the analogue photographic process	
multispectral acquisition	
Swath width larger compared with the traditional systems, considering same geometric resolution	Square or rectangular scenes acquisition as in traditional photogrammetry
Possibility of positioning of several sensors in the same focal plane allowing the multispectral and stereoscopic acquisition	Operative simplicity due to the analogy with the traditional photogrammetry
Larger sensors	Better geometric accuracy
Larger pixel	Simplex construction
Better signal-to-noise ratio	
Weakness	
Geometric resolution is still better using large format photographic films (23×23 cm^2)	
The multispectral acquisition by CCD are not perfectly registered	
Each multispectral band is not acquired from the same point of view	
The CCD can register data only in monochromatic mode	
Necessity of a DEM with high quality and definition to register the spectral bands	
The platform instability requires the presence of monitoring systems	Impossibility to design a matrix camera covering 23×23 cm^2
Necessity of more accurate geometric monitoring of the acquisition (GPS/INS)	Necessity to use combined cameras
Acquisition of *n* image centres to be registered	High quantity of collected data per single acquisition
Each scan line must be considered as a single acquisition to be registered with the following one	Difficulty in transmission of large amount of data between two acquisition
Long resampling process that reduce the final accuracy	Insufficient data transfer velocity (baud rate) in real-time acquisition
Necessity	
Accurate radiometric calibration of the sensors Appropriate store and conservation of the raw data Appropriate digital processing of the images after the acquisition Platform furnished by a laser scanning system for DEM of high quality Re-definition of the software programmes for the digital management system because not compatible with the existent hybrid systems of scanning and elaboration of analogue photos.	

Box 6.19 Comparison between parameters of three-linear and matrix digital acquisition in aerophotogrammetric systems

Parameter	Three-linear digital camera	Matrix camera
Sensor size	Linear CCD 100 mm	CDD area 60×60 mm^2
Number of pixel	10,000 (mono or three colours)	4000 × 4000 till 9000 × 9000
Pixel size	7–10 μ	8–12 μ
CCD technology	Advanced	In progress
Image geometry	Complex	Simple
Geometric calibration	Advanced	Advanced
Radiometric resolution	Excellent	Very Good
Colour	Easy acquisition	Sometime complex
Support systems	GPS/INS required	GPS/INS helpful
Data recording	Advanced	Advanced
Data compression	In progress	Advanced
Data transmission	Not limiting factor	Limiting factor

6.5.2 Hyperspectral Sensors

Hyperspectral sensors are imaging spectrometers with more than 16 bands. The high spectral resolution allows a detailed representation of the spectral response of each pixel. The main difference in comparison with the multispectral sensors is not in the sensor technology but in the methodologies for the digital data processing with an adequate bands selection in function of the applications for which they should be used. Table 6.22 reports the characteristics of the most used airborne hyperspectral sensors.

One of these, MIVIS hyperspectral sensor (*Multispectral Infrared and Visible Imaging Spectrometer*) is widely used in Italy and in Europe.

Table 6.22 Technical characteristics of some airborne hyperspectral sensors

Sensor	Number of bands	Spectral range (nm)	Spectral resolution (nm)	Pixel per line	IFOV (mrad)	Dynamic range (bit)
AVIRIS	220	0.40–2.45	9.4–16	550	1.0	12
CASI	15–288	0.43–0.95	2.9	512	1.02–1.53	12
GERIS	63	0.40–2.50	25–120	512/1024	2.5	16
HIRIS	192	0.40–2.45	10	2	30 m	12
MIVIS	102	0.43–12.7	8–500	765	2.0	12
MAIS	71	0.45–12.2	20–800	Variable		8
HYDICE	206	0.4–2.5	7.6–14.9	320	0.5	12
HYMAP	128	0.44–2.94	12–16	512	2.5 × 2.0	12
MODIS	36	0.42–14.5	10	1024	1.4	10–12

In order to provide the scientific community with a support activity to environmental research, the Italian National Council of Research CNR, within the climate, environment and land strategic programme for Southern Italy, launched the LARA project (*Aerial Laboratory for Environmental Research*) for the acquisition of remotely sensed data from aerial platforms. The project activated the scanning system MIVIS (Daedalus AA5000), first of a generation of hyperspectral sensors, operating with high geometric and spectral resolution, on board the CASA C212 aircraft. MIDAS software system (*Multispectral Interactive Data Analysis System*) controls the management, storing and distribution of MIVIS (Marino, 1993).

The characteristics of MIVIS hyperspectral instrument are reported in Box 6.20.

Box 6.20 Technical characteristics of the airborne hyperspectral instrument MIVIS

Parameter	Characteristics
Spectral bands	102
Spectral range	In the range: 0.43–12.7 μm
Reference black bodies	Number 2 in the range: − 15°C + 45°C
Spatial registration of the bands for pixel alignment	Optical Field Stop (IFOV 2.0 mrad)
Digital Field of view (FOV)	71.1°
Scan rotational velocity	Range: 6.25–25 scan/s
Components	Optical–mechanical with thermal compensation
Scan line	755 pixel
Data quality	Real-time computer monitoring
Dynamic range (radiometric resolution)	12 bit per pixel max scene temperature, without signal saturation, 1.200°C
Recording system	VLDS (Very Large Data Store); recording speed 6.2 Gigapixel per hour
Position (accuracy 12–40 m) and velocity (accuracy 0.05–0.20 m/s) of the aerial platform	PAS system (Position and Attitude Sensor) with GPS
Roll and pitch (accuracy 0.2°)	Gyroscope
Yaw control system of the aerial platform	
Operator interface	Touch screen display
Monitoring of two bands selected during the acquisition	Moving window display and digital oscilloscope

MIVIS is a modular instrument made by four spectrometers, with 102 bands simultaneously recording in the visible, near, medium and thermal infrared (Fig. 6.31).

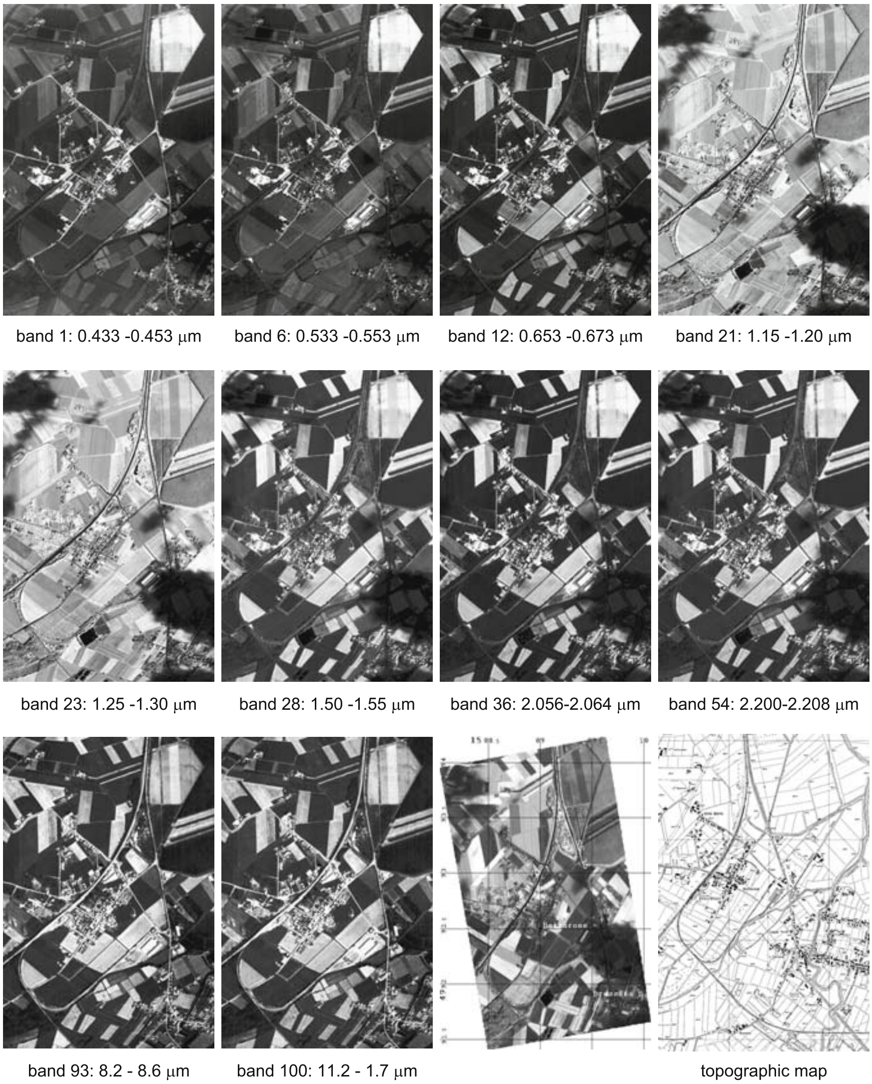


Fig. 6.31 Some spectral bands of the MIVIS hyperspectral radiometers in the visible, near, medium and far infrared originally acquired with 4 m geometric resolution (CNR-LARA, Italy)

The choice of the bands reflects the needs of applied research in scientific domains using remote sensing, such as botany, agronomic science, geology, pedology, hydrology, oceanography, atmospheric science (Table 6.23).

Sensor geometric resolution is 2 mrad, in function of the platform altitude: at 12,000 m the ground pixel resolution is 24 m × 24 m, the resolution progressively increases as the acquisition altitude decreases, until to 2 m × 2 m at 1000 m altitude.

Table 6.23 Spectral intervals of the four radiometers of the airborne hyperspectral instrument MIVIS and fields of application

Spectral bands	Spectrometer 1 (μm)	Field of application
1	0.433–0.453	Chlorophyll in water
2	0.453–0.473	Yellow substance in water
3	0.473–0.493	Pollution monitoring in water
4	0.493–0.513	Geological studies
5	0.513–0.533	Suspended material in water and monitoring
6	0.533–0.553	Coastal monitoring
7	0.553–0.573	Organic matter
8	0.573–0.593	Forestry, phytoplankton
9	0.593–0.613	Land Use/Land Cover, chlorophyll in water
10	0.613–0.633	Mineralogy
11	0.633–0.653	Geological maps
12	0.653–0.673	Geological maps
13	0.673–0.693	Tectonic, soil productivity
14	0.693–0.713	Acid rainfall
15	0.713–0.733	Erosion
16	0.733–0.753	Suspended material in water
17	0.753–0.773	Geological maps
18	0.773–0.793	Forestry inventory
19	0.793–0.813	Vegetation structure
20	0.813–0.833	Vegetation structure
	Spectrometer 2 (μm)	
21	1.150–1.200	Ferrous content in the soil
22	1.200–1.250	Ferrous content in the soil
23	1.250–1.300	Oxygen adsorption in the soil
24	1.300–1.350	Oxygen adsorption in the soil
25	1.350–1.400	Soil moisture
26	1.400–1.450	Soil moisture
27	1.450–1.500	Soil moisture
28	1.500–1.550	Soil moisture
	Spectrometer 3 (μm)	
29–92	2.000–2.500	Geological maps, mineralogy
	(64 bands, one every 0.008 μm)	Lithology, soil moisture
	Spectrometer 4 (μm)	
93	8.200–8.600	Thermal pollution monitoring
94	8.600–9.000	Chlorophyll in water
95	9.000–9.400	Fire maps
96	9.400–9.800	Geological maps
97	9.800–10.20	Soil moisture
98	10.20–10.70	Suspended material in water, water vapour
99	10.70–11.20	Water temperature
100	11.20–11.70	Oil spill
101	11.70–12.20	Lithology
102	12.20–12.70	Acid rainfall

The MIVIS is constituted by five different sub-system components:

- *Scanning head and spectrometer*: it is made by optical elements (primary collimating telescope and scanning rotating mirror), by a motor-counter, controlling the scanning, by two thermal reference black bodies and a structure containing all the components and providing the interface with the spectrometer.
- *Digitizer*: its main function is to convert the 102 analogical signals in 12 bit dynamic range, formatting and recording data on a magnetic support. Further functions include the collection of information about each scanning line and their insertion in the data recording.
- *Power distributor*: it acts as power converter and then controls the possible sources of electrical problems.
- *Moving Window Display and Monitor*: receives the video images from the digitizer for the simultaneous control of the two selected bands during the acquisition.
- *VLDS (Very Large Data Store) Recorder*: a digital recorder provided with a buffer enabling, through internal memories, the transfer of data recorded at different scanning speed. The recorded data can be later processed, stored and distributed by MIDAS analysis system.

6.5.3 Unmanned Aerial Vehicles

The *Unmanned Aerial Vehicles* (UAV) are a class of aircraft that can fly without the onboard presence of the pilot. They can be flown by electronic equipment present on the vehicle and at a *Ground Control Station* (GCS), or remotely piloted and operated by radio-controlled devices (*Remotely Piloted Vehicle*, RPV). In literature terms like Drone, *Remotely Operated Aircraft* (ROA), and *Unmanned Vehicle System* (UVS) are adopted to indicate such a category of vehicle.

The development of UAVs started in the 1950s for military purposes aimed at producing vehicles able to carry out missions without onboard pilots. The constant increase of research programmes and the evolving technologies have allowed implementing systems able to fully satisfy the requested requirements.

The success in the military field offers a valid stimulus for the development of such systems for civil applications, especially in the field of environmental monitoring, agriculture and public security.

The use of unmanned vehicles would ensure the possibility of carrying out long missions, without risking human losses and performing large-scale monitoring operations. UAV association classifies UAVs into three categories with respect to their possible usage. Each typology of aerial vehicle is subdivided according to their features and performance; particular reference is made to the vehicle's range, maximum climb rate, endurance and weight.

Table 6.24 refers to the *Tactic* group UAV systems, which nowadays are largely used for photogrammetric applications. The technological development of navigation systems enables the reduction of costs and the miniaturization of payloads. It is afterwards possible to set up projects based on low-cost platforms.

Table 6.24 UAV classification (UAV Association)

UAV categories (by UAV Association)	Acronym	Range (km)	Climb rate (m)	Endurance (h)	Mass (kg)
Tactic					
Micro	m (Micro)	< 10	250	1	<5
Mini	Mini	<10	150–300	<2	150
Close range	CR	10 a 30	3000	2–4	150
Short range	SR	30 a 70	3000	3–6	200
Medium range	MR	70 a 200	5000	6–10	1250
Medium range endurance	MRE	> 500	8000	10–18	1250
Low altitude deep penetration	LADP	> 250	50–9000	0.5–1	350
Low altitude long endurance	LALE	> 500	3000	>24	< 30
Medium altitude long endurance	MALE	> 500	14,000	24–48	1500

6.5.3.1 Tactical Group, the Mini-UAVs

If satellite or aerial aerophotogrammetric data are not available or not suitable to supply radiometric and geometric information, in situ missions must be foreseen. The fields of interest are very wide, but the most concrete application is the creation of basic and thematic maps for emergency management in remote areas. For example, ITHACA Association (Bendea et al., 2007) has developed a low-cost mini-UAV. *Pelican* is a remote-controlled aircraft equipped with a GPS/IMU navigation system and different photographic sensors suitable for digital photogrammetric shootings with satisfying geometric and radiometric quality. It can be easily transportable on normal aircraft and is usable on the field by a couple of operators.

The project carries on operational and research activities in geomatics for analysis, evaluation and mitigation of natural and man-made hazards, in supporting the capacity of the *United Nations-World Food Programme* (UN-WFP) to prepare for and respond to humanitarian emergencies.

The mini-UAV prototype (MH2000, with reference to the wingspan in millimetres) is a remote-controlled aerial platform equipped with a GPS/IMU navigation system and digital sensors devoted to photogrammetric surveys (Fig. 6.32). It can be easily transportable on normal aircraft and is usable on the field by a couple of operators. The aerial platform is developed and patented by the *department of Aerospace Engineering* (DIASP) of the Politecnico di Torino, Italy. The MH2000 platform is characterized by a conventional configuration (Table 6.25): fixed wing, tailless integrated wing-fuselage, tractor propeller driven, all electrically powered. The fixed-wing configuration gives the aerial vehicle a better capability of withstanding adverse weather conditions, such as gusts, and it allows larger payload capabilities and a superior flight performance.

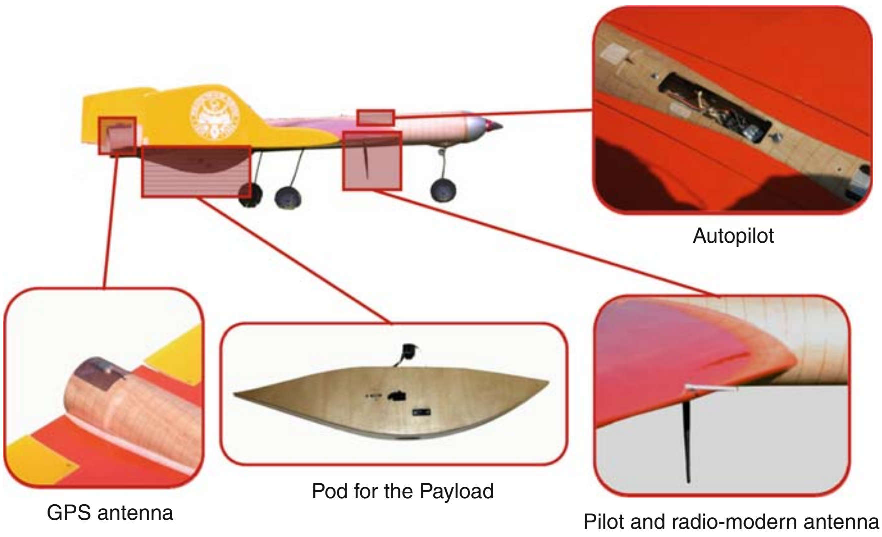


Fig. 6.32 The UAV prototype MH2000

The aircraft is equipped with the Micropilot MP2128g which is an autopilot assuring autonomous flights and providing a real-time attitude of flight. It includes a GPS unit, three-axis gyroscope and accelerometer (IMU), relative airspeed probe, pressure altitude transducer, AGL ultrasonic altitude sensor, and external servo board.

Table 6.25 MH2000 technical specifications

Specifications	MH2000
Wing span	2000 mm
Wing surface	2.1 m2
Length (fuselage)	1750 mm
Width (fuselage)	150 mm
Payload weight	6.7 kg/m2
Weight (body)	10 kg
Max take-off weight	14 kg
Payload	2.5 kg (25%)
Lithium-ion batteries	1.0 kg (10%)
Motor	0.5 kg (5%)
Motor model	Hacker C50-21 XL
Output power	2 hp @ 6600 rpm
Propeller	APC 16 in.$\times$ 8 in.
Min speed	8 m/s (29 km/h)
Max speed	20 m/s (72 km/h)
Optimum range limits	25 km
Optimum endurance	1 h

The UAV configuration allows carrying onboard digital sensors for video and imagery acquisition. The sensors are mounted in a pod (Fig. 6.32) placed on the belly of the fuselage, which can be remotely controlled by the Ground Control Station (GCS) through a direct link with the autopilot board. A real-time video downlink of the over-passed areas is made possible exploiting the video output signal of the camera. At a flying altitude of 200 m and aircraft speed of 15 m/s, images characterized by a Ground Sample Distance (GSD) of 75 mm with 257 m side coverage are acquired. The shooting time interval required for 60% along-track overlap for stereoscopic view is about 4.5 s.

6.5.4 Laser

The Laser is an applicative device that transforms the primary energy (electrical, optical, chemical, thermal or nuclear) in a monochromatic and coherent beam of electromagnetic radiations of high intensity: the *Laser light*. Albert Einstein is the father of the fundamental discover of the laser light in 1917.

Generally the acronym LIDAR (*Light Detection and Ranging*) is used to refer to environmental remote sensing systems based on the use of laser (*Light Amplification by Stimulated Emission of Radiation*).

The LIDAR aims to individuate and determine some characteristics and the measurement of the distance of an object by the emission of electromagnetic beams reflected by the irradiated object.

LIDAR instruments' project logic is similar to radars, but with a significant difference in the type of emitted radiations: in the case of LIDAR the laser transmitter can generate a luminous beam, collimated by an adequate lens and mirrors system, in an interval of the electromagnetic spectrum typical of the optical frequencies (0.3–15 μm); the LIDAR is not a scanning system, it does not produce images but punctual measures and is generally used for studies about terrestrial atmosphere components. Whereas the laser scanner, used for studies of the terrestrial surface, produces images and, as an active system, in theory can be used 24 h a day.

The laser system must not be confused with the radar (see Chapter 4) active system operating in the range of microwaves (0.1–100 cm).

The principle of laser technology, when adapted in remote sensing systems, retrieves tri-dimensional numerical maps and it performs detailed topographic measures with more precision and in a shorter time than traditional methods. The laser scanning system, or laser scanner or laser altimeter, is an instrument that emits a signal and records its backscattered radiation measuring with precision the distance between the platform carrying the laser and the irradiated surface.

The *Aerial Laser Scanning* (ALS) is the convergence of three technologies:

- *Laser rangefinding*;
- *Satellite Positioning System* (GPS, GLONASS, Galileo);
- *Inertial Navigation Systems* (INS).

The system's peculiarity is to also carry out measurements in high vegetation density and to characterize position, shape and height of different objects and infrastructures. The problem of the shade of the irradiated objects does not exist, but the occlusion due to the presence of other objects between the sensor and the object is decisive.

The reflection of many objects is high in the monochromatic range typical of the laser and thus causes sensor saturation. In some cases some objects can be invisible due to their low reflectance in the laser spectral region obtaining a better penetration (Baltsavias 1999a, Krauss & Pfeifer 1998).

The method is based on the measure of distances between an airplane, or a helicopter, and the terrestrial surface measuring with precision the time taken by the laser signal's pulse to reach the ground and, reflected, to go back to the emitting platform (Fig. 6.33).

The reflected laser signal is received by a small telescope collecting the laser light in a detector. The time taken by the signal is converted in distance from the platform considering the laser beam speed equal to the light speed in the vacuum (2.99×10^{10} cm $\cdot$ s^{-1}).

The wavelength of the laser signal transmission, in relation to the instrument, is included within the spectral range 0.2–1.5 μm of the electromagnetic radiation in ultraviolet (0.2–0.4 μm), visible (0.4–0.7 μm) and near infrared (0.7–1.5 μm) the first and the last external to the human eye's visible field (Table 6.26).

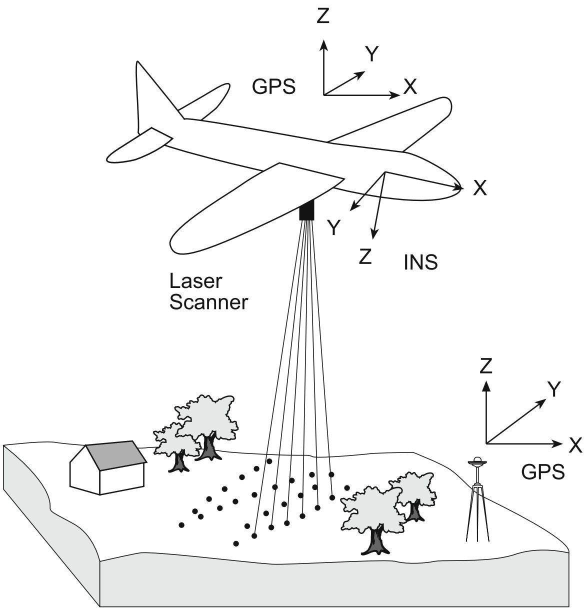


Fig. 6.33 The laser scanner system

Table 6.26 Main characteristics of some Laser scanner systems

System	*ALTM 1020	TopoSys	TopEye	**ATM II(NASA)	ScaLARS
Measurement method	Propagation time of a radar pulse				Phase difference measurement multi-frequency
Spatial information	3D points				3D points and reflectance values
Laser wavelength	1.047 μm	1.535 μm	1.064 μm	0.523 μm	1.064 μm
Acquisition altitude	1000 m	1000 m	500 m	700 m	750 m
Resolution	0.03 m	0.10 m	0.10 m	0.10 m	0.06 m
Frequency	2000 Hz	80 kHz	6000 Hz	2 a 10 kHz	7000 Hz
Swath width	0.25 – 1000 m	0.20 – 1000 m	0.60– 500 m	0.60– 700 m	1.50 ÷ 750 m
Scan system	Floating mirror	Rotating mirror and optical fibre	na	Rotating mirror	Rotating mirror
Scan frequency	50 Hz	600 Hz	na	20 Hz	na
Scan angle	Variable till ± 20°	± 7°	Variable till ± 10°	± 15°	± 14° / ± 20°
Points density	1 point per 2.6 m^2	5 points per m^2	1 point per 4 m^2	1–4 points per m^2	1 point per 4 m^2

na: not available; *ALTM: Airborne Lidar Topographic Mapping; **Airborn Terrain Mapper

The laser is characterized by:

- monocromaticity: signal composed by only one frequency of light;
- spatial coherence or unidirectionality: the signal moves in the space in a defined direction;
- temporal coherence: the signal is constituted by waves of the same frequency and same phase generating a pulse of high intensity and power.

The laser transmitters emit from 2000 to 33,000 pulses per second; thousands of distances equal to the number of pulses emitted by the transmitter over a very narrow strip are recorded by scanning the investigated area through a rotating mirror.

The measured distances are converted in planimetric coordinates and altimetric values for each laser pulse combining the measures of the distance with the exact position of the instrument on the acquisition platform. The signal emission's aerial point along the flight line is identified with precision by a satellite positioning system (GPS, GLONASS and Galileo) applying the technique known as kinematics differential solution. Through the differential GPS, a few centimetres accuracy can be obtained. A reference GPS ground station collects and records data later processed together with the data collected by the onboard GPS.

In order to correctly identify the 3D coordinates of the laser signal, the pulse orientation must be known with precision. The aerial platform is subject to a possible non-constant advancing speed as well as roll, pitch and yaw errors, referring to rotations on the respective axes.

The laser pulse direction is defined with precision using an *Inertial Navigation System* (INS) which records these errors and measures the platform orientation. The Inertial Systems can detect errors smaller than 0.01°.

The measure of the distance and the defined signal emission – reception angle allow determining of the spatial coordinates of the reflecting point to be localized. Combining the collected information about distance, position and direction, each laser signal can be georeferenced.

Acquiring in quick succession parallel lines orthogonal to the fly direction of the aerial platform, continuously led by an exact positioning system, it is possible to map wide areas, with side overlap for swathes mosaicing.

Data from the single acquired swathes are processed to produce 3D terrain maps, as Digital Surface Models (DSM) and Digital Elevation Models (DEM) describing the topography using a regular grid with altimetric values. This method is alternative to the acquisition of contour lines from stereo-pairs of aerial photos by photogrammetry techniques, satellite stereoscopy and/or interferometric methods.

With the most recent laser scanning systems models, it is possible to measure multiple reflections for each pulse: for example, when detecting a wood the first reflection is from the top of the trees, the second from the branches and leaves and the third from the ground, as the radiation penetrates in the canopy. With these different reflections at different levels, it is possible to obtain several profiles: from the tree top to the ground surface (Fig. 6.34). This system records up to five reflections from one pulse. The last reflection, at ground level, reproduces the DTM. According

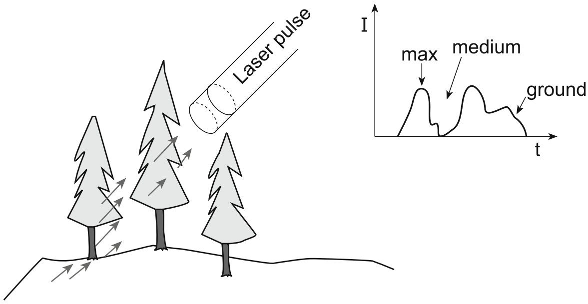


Fig. 6.34 The flight altitude influences the ground resolution. In woody areas, several backscattered signals (eco) can be recorded at different heights in the vegetation. The correct separation of the points belonging at the bare soil and at the top of the vegetation permits the building of the Digital Terrain Model (DTM) and/or the Digital Surface Models (DSM). I: Intensity, t: time

to its generation process, the best acquisition occurs when plants are without leaves (Turner, 2000).

The recorded signals can be used to measure the biomass height based on the difference between the tree top and ground measures. An identical process is followed in measuring infrastructures and buildings' height. The possibility of clearing the effective terrain morphology from the vegetation cover defines with precision, for instance, flooded areas also in the presence of high vegetation density. The altimetry measurements' accuracy can restrict the error at ± 15 cm (Sapeta, 2000).

The applications include hydrogeological risk, river and flooding areas, hydrological models, urban planning, identification of paths for oil pipelines, etc.

The acquisition of laser scanner can contain more than 100,000 points/km^2.

6.5.4.1 Laser Scanning System

The limiting factor for the geometric positioning accuracy of the *Airborne Terrain Mapper* (ATM II), NASA's laser scanner, for instance, is the capacity of the Inertial Navigation System (INS) in correcting the movements from the ideal flight line of the platform. The most accurate in terms of geometric positioning is the laser scanning system, which can give precision of the order of 5 cm. The positioning systems' resolutions range between INS and laser. The combination of the resolutions and accuracy of the three sub-systems determines LIDAR instrument final resolution.

Aerial laser systems operate within the near infrared in the spectral range 1040–1550 nm (1.04–1.55 μm). The pulses are one close to the other, with a medium-high power frequency ranging from 2 to 80 kHz. The beam divergence is about 1 mrad. The swath width is a function of the platform altitude and scanning angle. In general the field of view (FOV) is lower than 30°, so the swath width ranges from 100 to 750 m for a flight height from 200 to 1500 m. The flight height influences the acquisition accuracy: with higher altitudes the angular errors increase, as the points do not form

a regular grid on the ground surface and the concept of points is quite vague due to the beam divergence.

The laser scanner is an electro-optical device based on the principle of optical triangulation. Commercial laser scanners are often calibrated in the scanning plane to account for variation of the incident angle of the laser beam.

The beam projection area can be many square metres wide for surveys from higher altitude, decreasing the geometric and radiometric quality (Fig. 6.34).

Other factors which can influence the acquisition quality are strong winds, snow, rain, fog, humidity and low clouds.

6.5.4.2 Airborne Laser Scanning (ALS) Acquisition

In order to properly plan an altimetric acquisition with an airborne scanning system, it is necessary to collect GCPs properly distributed over the area of interest. Large flat areas can be used for instrumentation test.

The risks of failing during the laser scanning data acquisition process are quite high, depending on the accuracy in considering many factors. If the acquisition is successful, the collected data are of very high value as georeferenced in a coordinate system.

An example of point's localization error is reported in Fig. 6.35. Assuming

- laser localization (X_o,Y_o,Z_o) is error free;
- medium flight height equal to 500 m;
- Gaussian distribution of the INS errors for the angles $\angle\, a$ and $\angle\, b$, with standard deviation of 0.1°, equal to 0.001745 rad,

the error in the laser reading interval r is expected to have normal distribution with average equal to 0 and standard deviation of the flight height equal to 0.1%.
The processing operations of the collected laser data include:

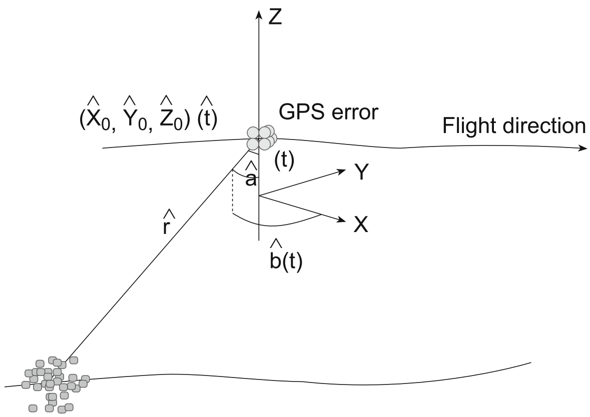


Fig. 6.35 The errors of the GPS/INS system are at the base of the inaccuracy of the localization of the measured points

- automatic reflectors or geometric entities extraction;
- points cloud filtering (DDSM: Dense Digital Surface Model) for the elimination of outliers, gross errors and of basic noise;
- discharge of points not representatives of the observed object of interest, as background and close up;
- automatic alignment of adjacent tri-dimensional models;
- geo-coding of the clouds of points in an external reference system a priori known;
- laser scanner bundle block adjustment for a correct alignment of a series of adjacent—acquisitions;
- image external orientation parameters re-adjustment (if the laser is integrated with digital camera) or photogrammetric image survey and processing (in case of laser not integrated with digital camera);
- colour assignment to the points cloud;
- image mapping.

The result is a complex and complete cloud of points of the observed object representing the correct starting point for successive processing and mapping.
The final products can be of two types:

- products obtained only from the laser technology as 3-D representation, contour lines, sections, aspect models, etc.;
- products obtained by integration of laser scanner technology with photogrammetric techniques as precision ortophotos, solid images, 3-D colour models, etc.

A *solid image* can be obtained by a digital procedure that permits maintaining the original geometric and radiometric characteristics by photogrammetric acquisition of the digital image with possibility of elaboration of the tri-dimensional data collected by the laser scanner system.

6.6 Instruments for In-Field Acquisition

Instrumental ground observation permits non-destructive and proximal sensing. Using traditional photographic instruments, digital cameras, video cameras in the visible and near infrared, thermal cameras for studies in the thermal bands, radiometers and spectroradiometers, it is possible to obtain images and measures of the objects and study their condition in real or near real time.

6.6.1 Photographic Films

The reflection of the solar radiation by the vegetation in the wavelength interval including the visible and the near infrared (NIR) depends on two parameters: pigment content and leaf internal structure. Actions able to modify one or both the parameters have as a result the variation of the vegetation spectral response.

In the infrared (IR) colour film, eliminating the blue light with a proper filter (*Wratten 12*), the three layers are sensitive to the green, red and infrared radiations. As part of the green is reflected by the surface, while the IR is reflected by the surface and scattered by the deepest tissues, the panchromatic appears clear, whereas the IR is more confused. The IR false-colour composite highlights a decrease in vegetation reflection; in fact unhealthy crops gradually lose their capacity of reflecting in the near IR. Such an experience can be carried out by measuring the density of the three sensitive layers of IR-colour aerial acquisition films, over portions of fields with different crops and diseases, as well as using IR-colour photos acquired from the ground, fastened balloons, helicopters and UAV.

Photographic acquisitions from very low altitude, or in-field acquisitions, allow highlighting of the anomalies of spectral behaviour in the near infrared in relation to visible phenomena. A limit of these techniques is related to the impossibility of evaluating in real time the acquisitions, as it is necessary to process, print and scan them.

Digital photographic cameras allow this limit to be overcome by different tools sensitive to the three intervals of the visible (0.4–0.7 μm), near IR (0.7–0.9 μm) and medium IR (3.0–5.0 μm). The amateur reflex digital photographic camera is based on the CCD (*Charge-Coupled Device*) sensors technology and acquires images of 24 mm $\times$ 36 mm surface, equal to common photographic films. This unprofessional camera can acquire matrix of 5 million pixels (1800 $\times$ 2700) and more.

The thermal digital camera, with a CCD sensor, acquires matrix of 500 $\times$ 800 pixels in the interval 3.0–5.0 μm, with an internal cooling system that does not need the coolant to be replaced. The images can be instantaneously transferred to a PC and processed.

6.6.2 Digital Photo Camera

The main characteristics of digital photo cameras sensors are similar to the photogrammetric cameras and can be identified in three types of acquisition systems (see above):

- central: *matrix* systems;
- non-central: *punctual* or *linear* scanning systems.

The time of exposition of a photo camera with a matrix sensor is generally short-medium from 1/30 to 1/200 of a second. Very short exposition like 1/8000 of second is possible in some models.

The CCD can acquire only in monochromatic mode. This limit is overcome by artifices, such as

- single shot with three CCD: the light, after passing through the lens, is split, for example by a Newton prism; three different sensors record the information of red (R), green (G) and blue (B), and the image is reproduced in RGB colour image;

- single shot with one CCD: each pixel is covered by linearly or mosaic arranged RGB micro-filters;
- triple shot with only one CCD: coloured filters rotate in front of the matrix sensor with the acquisition of the three main colours in three separate exposures later recombined in the final image. This acquisition can be adopted if the recorded scene or objects are stationary.

An innovative acquisition method introduces a fourth cyan colour filter (*Emerald*) that, with the traditional red, green and blue, makes the so-called RGBE filter (*Red, Green, Blue, Emerald*). The reason for adding the fourth filter colour is to reduce colour reproduction errors and to record natural images closer to the natural sight perception of the human eye (Plate 6.9). This method adopts the standards defined by the ITU (*International Telecommunication Union*) and by the IEC (*International Electrotechnical Commission*) that specify the relation between non-linear RGB values and the equivalent CIE system's X, Y, Z through transformation matrices into the system called sRGB (*standardized Red, Green, Blue*); the *International Colour Consortium* (ICC) has then defined the colour profiles to which they refer.

The colour definition depends on the acquisition and reproduction systems; this problem is overcome by the sRGB system with a univocal reference to the original colours. The RGBE filter adopts the sYCC system transformed in irreversible way into the sRGB system.

Y, Cr and Cb derive from R, G and B this way:

$$
\begin{aligned}
\mathrm{Y} &= 0.299\mathrm{R} + 0.587\mathrm{G} + 0.114\mathrm{B}\ \text{(Luminance or B/W signal)}\\
\mathrm{Cb} &= 0.564(\mathrm{B} - \mathrm{Y}) + 1/2\ \text{full scale}\ \ \text{(a colour difference signal)}\\
\mathrm{Cr} &= 0.713(\mathrm{R} - \mathrm{Y}) + 1/2\ \text{full scale}\ \ \text{(the other colour difference signal)}
\end{aligned}
$$

Digital photo cameras' geometric resolution can be largely variable: with 300 dots per inch (dpi) a low level model acquires 640×480 pixel image, equal to $5 \times 3.8\ \text{cm}^2$. A high-level model, with 300 dpi, can record an image area larger than 6000×7500 pixels, corresponding to $50 \times 65\ \text{cm}^2$.

The image quality depends on the *radiometric depth*, or dynamic range, measured in bit, that is the amount of data the CCD can record (see Chapter 4).

Practical models are usually 8 bit; the professional can reach 12 bit (4096 level) and more. A normal photographic film can distinguish 32 tonal levels corresponding to 5 bit resolution.

The digital image radiometric resolution depends on the analogical–digital converter. If the sensor can record 8 bit dynamic range, the analogical – digital converter should have a higher resolution, i.e. 10 bit, to be sure it contains all the values.

The software most commonly used for both the conversion of the CCD sensor signal and for the image print generally has 8 bit resolution per colour. A higher radiometric depth, i.e. 12 bit, allows a better detail in the highlights and shades, but if the signal is converted from analogical into digital with 10 bit per colour and the

printer or the plotter can manage only 8 bit per colour, the additional information from 8 to 12 bit is lost.

Maximum quality can be obtained using potentiality of the digital photo camera. It is possible to use compression techniques to reduce the image file size increasing the number of images stored in the memory. However, increasing the compression, the quality could be lower if the compression modes and factors are not controlled.

6.6.3 Terrestrial Laser Scanner

Terrestrial Laser Scanning (TLS) is a significant technique for the collection of 3D geospatial data in remote survey practice, analysis and research. TLS offers an alternative to traditional survey techniques. It consists of automated high-speed data capture of complex surfaces and often inaccessible environments. Amongst innumerable variety of applications, terrestrial laser scanning is a valuable tool in the areas of survey engineering, mine engineering, architecture and heritage management.

An object has to be scanned from different view points. Afterwards, the aim is to register single point clouds to one common point cloud. Single scans from different scan positions are registered to a local coordinate frame defined by the instrument. For data processing, the scans must be transformed into a common coordinate frame. This operation is termed as registration. For registration, homologous points (tie points) or objects (e.g. spheres, cylinders) are required.

Due to this drawback, automatic matching methods for multiple scans are needed and developed in specific software programmes. The intention is to have fast, robust and quasi-automatic methods for registration, which can be carried out in the field as far as possible.

Terrestrial laser scanners can be classified on the conceivable applications. There is no one universal laser scanner and the choice depends on the application. The first difference concerns with indoor use and medium ranges (up to max. 100 m) and outdoor use with long ranges (up to several 100 m). Terrestrial laser scanners can be categorized by the principle of the distance measurement system, correlated both to the range and to the accuracy. Most of the laser scanners are based on the time of flight principle. This technique allows measurements of distances up to several hundred of metres. The advantage of long ranges implies a less accuracy in the distance measurement (approximately 1 cm). The phase measurement principle represents the other common technique for medium ranges. The range is restricted to 100 m. In opposite to the time of flight principle, accuracies of the measured distances within some millimetres are possible. Close range laser scanners with ranges up to few metres are also available. However, these scanners do not fit into the classification of *terrestrial* laser scanners. They are used for industry applications and reverse engineering (Table 6.27).

Table 6.27 Classification of terrestrial laser scanners

Measurement system	Range (m)	Range accuracy (mm)
Time of flight	~ 1000	> 10
Phase measurement	< 100	< 10
Optical triangulation, laser radar	< 10	< 1

6.6.4 Video Cameras and Thermal Cameras

The testing and the use of ground video cameras in the visible and infrared have applications mainly in the laboratory, due to considerable problems in normalizing both low-altitude aircraft and ground acquisitions. With the availability of cheaper and quicker hardware/software instruments (GPS, INS, etc.), the video technique acquisitions are more practical and suitable for real-time monitoring.

Video cameras, connected to a processor, acquire digital scenes that can be processed as for the satellite and/or aircraft imagery. Spectral bands can be selected with the use of appropriate filters, as for photographic acquisitions. Lenses are interchangeable, allowing configuration of different focal lengths. Video cameras can operate in the range of the visible or, according to the sensors' technical characteristics, in the near-medium infrared up to 2.0 μm. Video cameras operating in the visible and infrared intervals are suitable for studying the spectral behaviour of healthy or stressed vegetation.

Thermal cameras are used to study and measure the black body surface temperature of the objects. The spectral intervals range from 3.0 to 5.0 μm and from 10.5 to 12.5 μm in which the energy emitted by the investigated objects can be detected. The thermometric analyses require more complex image processing and high experience in interpretation.

Thermal cameras are suitable for measuring thermal inertia of objects. Thermal inertia is a measure of the thermal mass and the velocity of the thermal wave which controls the surface temperature of a matter. In field investigations errors in measurement can be generated by thermal variations near the investigated object, often higher than the thermal resolution of the instruments (0.1–0.5°C). Thermal cameras are also used for medical applications.

6.6.5 Radiometers

Radiometers are portable electro-optical instruments for measuring the wavelength and the radiant energy emitted and/or reflected from the surfaces. Two kinds of radiometers can be distinguished:

- *single band*, in which radiance measurements are carried out in single spectral bands corresponding to the ones installed on the satellites;

- *continuous band*, or *spectroradiometers*, in which the measurement is simultaneously in the range from 0.35 to 3.0 μm by several narrow spectral bands (<0.008 μm) one subsequent to the other enabling in continuum the spectral behaviour of an object.

The radiometer is a device that measures the intensity of the incident radiation on the detectors. According to the interval of the spectrum, they are given different names:

– photometers, sensitive to visible and near infrared light;
– near, medium or thermal infrared radiometers;
– multispectral radiometers, operating in more bands;
– microwave radiometers, if operate in the spectral range from 1 to 100 m.

In general radiometers are made by an element collecting the radiations (optical antenna), a sensitive element or detector, a reference sample source and an instrument visualizing the measure. They are used both for in-field measurement and for atmosphere transparency assessment at the moment of the acquisition.

For example with the four bands of the radiometer Exotech 100 AX, it is possible to carry out measures of the radiance by interchanging optics corresponding to the spectral bands in the visible and near infrared of satellite sensors (e.g. ETM+, MERIS, ASTER). From the radiance values spectral reflectance values are retrieved.

Its weight is about 2 kg, set of optics and batteries included; thanks to this the instrument is easy to carry and to be used in the field. If fixed on a 3 m high pole perpendicular to the ground, the radiometer, using optics with 20° field of view, can measure the radiance of a circular area of about 0.5 m2.

The standard procedure schedules first the measurement of the sky irradiance, far from objects that can produce reflections on the instrument, which would invalidate the measures. Then the radiance of the object is measured over homogeneous sample areas. In order to obtain the measurement of reflectance, the ratio between the radiance of the vegetation and the incident radiation values (reflectivity or albedo) is calculated.

6.6.5.1 Photometers

The multispectral photometer is an instrument conceived by Volz (1959) to obtain direct measurement ($W{\cdot}m^{-2}$) of irradiance in narrow and centred wavelengths in visible and near infrared spectral ranges. These intervals must be chosen so that the radiation is not considerably absorbed by the atmospheric gases like oxygen, water vapour and ozone, but mostly subject to scattering attenuation.

Pointed towards the Sun, the photometer provides as many signals as the spectral intervals defined by the interferential filters in the instrument. The photometer retrieves values of the atmospheric optical depth at different wavelengths.

The signal measurement, proportional to the incident monochromatic solar irradiance, proceeds according to the following steps:

- pointing the photometer, set on a tripod, towards the Sun, centring its shape with accuracy and precision;
- recording the Sun temperature measure;
- selecting in succession on the photometer the interferential filters corresponding to the desired wavelengths (λ);
- when the value is stabilized, value recording, corresponding to the direct irradiance ($W \cdot m^{-2}$);
- apply the solar photometry method for measuring the atmospheric optical depth values to that λ .

This method refers to the Bouguer – Lambert – Beer's extinction law that relates the absorption of light to the properties of the material through which the light is passing.

6.6.5.2 Spectroradiometers

A spectrum-measuring radiometer is called a *spectroradiometer*. Currently some types of instruments immediately providing not only values of radiance and irradiance, but also those of reflectance, like spectroradiometers, are preferred to radiometers. The use of portable spectroradiometers makes possible in field high-resolution acquisitions of spectral responses and their comparison with reference curves recorded in the laboratory. The electromagnetic spectrum region covered by these instruments ranges in the interval 0.35–2.5 μm, i.e. from visible to near-medium infrared. The Field of View (FOV) can vary from 1.0° to 15° with the possibility of changing the optics in some models.

The spectroradiometers operating also in the medium infrared are heavy and costly.

6.7 Summary

The acquisition of imagery of the Earth's surface can be obtained using devices on satellite, airborne and/or ground platforms.

Images can be acquired as analogical or digital. Analogical photographic images are recorded in the visible, very near infrared interval of the electromagnetic spectrum. Digital images are recorded with different instruments from ultraviolet to microwave. The development of instruments for recording multispectral signals allowed the acquisition of quantitative information concerning the objects on the Earth. The acquisition geometry of the recording instruments can be with central perspective, optical–mechanical (i.e. Landsat sensors) or electronics elements in series along an array, as with the *whiskbroom* and *pushbroom* systems. The most used detector is the Charge-Coupled Device (CCD) that measures and records the

luminous energy, then uses the digital conversion process. The main instrument characteristics are defined by geometric, spectral, radiometric and temporal resolutions. The geometric resolution defines the *Instantaneous Field of View* (IFOV) and the *Field of View* (FOV).

Satellite platforms are meteorological, for weather forecasts and the study of the atmosphere, and/or for Earth Observation (EO). The age of the satellite for EO started in the 1960s, and the first artificial satellite *Earth Resources Technology Satellite* (ERTS) or Landsat-1 was launched on July 23, 1972. Satellite platforms planned for Earth Observation are in continuous development, and several international and government organizations work on the development of space programmes. One of the main organizations is the *Integrated Global Observing Strategy* (IGOS), which uses the main remote and ground sensing systems for global observation. Practical collaborations are within the *Group of Earth Observation* (GEO), with the task to build a worldwide Earth Observation programme: the *Global Earth Observation System of Systems* (GEOSS).

Several Space Agencies, both international and national, are developing programmes for studying, observing, monitoring and assessing terrestrial resources. The *Earth Observing System* (EOS) by NASA, for instance, commits in technological development of satellites, instruments and sensors and caring about data calibration and validation. EOS includes links and activities with other space agencies. The European Space Agency (ESA) has proposed the *Living Planet* programme, which will affect European and global investments and choices about space studies and applications for the next 25 years. The international cooperation is conducive for the realization of the most complex mission in history, the *International Space Station* (ISS). The range of research on ISS includes a wide spectrum of disciplines: biology, bio-medical, biotechnology, physics, fluid physics, material science, quantum physics, astronomy, cosmology, meteorology and Earth Observation.

The authorization for the property and the activity, in practice with no restrictions, of satellites for Earth Observation has stimulated competitiveness among the largest aerospace companies in the world to create new geospatial products and services required by the market. Several commercial satellites have been launched and possible applications include geospatial information, geo-marketing, real estates, insurances, etc. Commercial systems have geometric resolution from 0.5 m up in the panchromatic and from 2.0 up in at least three, or more, spectral bands. The planning of the missions generates a high frequency of acquisition reducing the revisiting period to 1–5 days and in some cases (EROS, COSMO/SkyMed) even more acquisition in the same day.

The airborne systems have larger flexibility, adaptable to every specific need, especially temporal. However airborne images have more distortions than satellite images mainly due to the instability of the acquisition platform. Aerophotogrammetric systems are described in Chapter 3. The evolution in this field concerns the development of digital photogrammetric cameras. There are two different acquisition methodologies: linear and matrix. The ADS40 and the HRSC-A are of linear type, and they solve the problems of stereoscopic acquisition and data transfer

velocity using three series of sensors set on linear arrays. DMC (*Digital Modular Camera*) bases its operative principle on the reconstruction of a central perspective, like in the traditional photogrammetric camera.

Hyperspectral sensors generally operate on spacecraft with more than 16 bands. The MIVIS system is reported as an example of this category of instrument.

A new group of applications includes the use of mini *Unmanned Aerial Vehicles* (UAV), aircraft that can fly without the onboard presence of the pilot. The use of unmanned vehicles would ensure the possibility of carrying out missions, without risking human losses and performing large-scale monitoring operations. The UAV configuration allows carrying onboard digital sensors for video and imagery acquisition. The fields of interest are very wide, but the most concrete application is the creation of basic and thematic maps for emergency management in remote areas.

The *Aerial Laser Scanning* (ALS) retrieves tri-dimensional numerical maps in shorter time than traditional methods. The ALS emits a signal and records its backscattered radiation measuring with precision the distance between the platform carrying the laser and the irradiated surface.

Instrumental ground observation permits non-destructive and proximal sensing. Using traditional photographic instruments, digital cameras, video cameras in the visible and near infrared, thermal cameras for studies in the thermal bands, radiometers and spectroradiometers, it is possible to obtain images and measurements of the objects and study their condition in real or near real time.

Further Reading

Borengasser M., Hungate W.S., Watkins R.L., 2008, Hyperspectral Remote Sensing, Principles and Applications. CRC Press Taylor & Francis Group, London, n.L1654, ISBN: 978-1-56670-654-4.

Chen C.H., 2007, Image Processing for Remote Sensing. CRC Press, Taylor & Francis Group, 400 pp, ISBN: 978-1 4200-6664-7.

ESA, 2008, The Living Planet Programme, new ed. European Space Agency.

Harding D.J., 2000, Principles of Airborne Laser Altimeter Terrain Mapping, NASA's Goddard Space Flight Center, Greenbelt, USA, March 17, 2000.

NASA, National Aeronautics Space Administration, 1993, EOS References Handbook. Washington DC, USA.

Paine D.D., Kiser J.D., 2004, Aerial Photography and Image Interpretation, 2nd ed. John Wiley & Sons, Inc., New York, 632 p.

Bibliography

Ackermann A., 1999, Airborne laser scanning – present status and future expectations. ISPRS Journal of Photogrammetry and Remote Sensing, 54(2–3): 64–67.

Almer A., Schnabel T., Raggam J., Gutjahr K., van Dahl M., 2007, Rapid information flow within a crisis management system. Proceedings of the 27th EARSeL Symposium, GeoInformation in Europe, M.A. Gomarasca (Ed.), Millpress, The Netherlands, pp. 455–462, ISBN 9789059660618.

Askey P., 2004, Sony Cybershot DSC-F828 Review.
Asrar G., 1985, Theory and Applications of Optical Remote Sensing. John Wiley & Sons, New York, USA.
Babington-Smith C., 1957, Air Spy: The Story of Photo Intelligence in World War II. Harper, New York.
Babington-Smith C., 1985,ir Spy: The Story of Photo Intelligence in World War II. American Society for Photogrammetry and Remote Sensing, Falls Church, Virginia (reprint).
Baltsavias E.P., 1999, A comparison between photogrammetry and laser scanning. ISPRS Journal of Photogrammetry and Remote Sensing, 54: 83–94.
Baltsavias E.P., 1999, Airborne laser scanning: basic relations and formulas. ISPRS Journal of Photogrammetry and Remote Sensing, 54(2–3): 199–214.
Bendea H., Chiabrando F., Giulio Tonolo F., Marenchino D., 2007, Mapping of archaeological areas using a low-cost UAV: the Augusta Bagiennorum Test Site. XXI International CIPA Symposium, 01–06 October 2007, Athens, Greece.
Bendea H., Boccardo P., Dequal S., Giulio Tonolo F., Marenchino D., 2007, New technologies for mobile mapping. 5th International Symposium on Mobile Mapping Technology MMT′07, Padua, Italy, 28–31 May 2007.
Bornaz L., Dequal S., 2004, The solid image: An easy and complete way to describe 3D objects. In: Volume XXXV part B5. XXth ISPRS congress. Istanbul. 12–23 July 2004, pp. 183–188. ISBN/ISSN: 1682–1777.
Burrows W., Deep Black E., 1986, Space Espionage and National Security. Random House, New York.
Colwell R.N., 1983, Manual of Remote Sensing (Vol. 1 and 2). American Society of Photogrammetry, Falls Church, Virginia, USA.
Eisenbeiss H., 2004, A mini unmanned aerial vehicle (UAV): system overview and image acquisition. International Workshop on Processing and Visualization Using High-Resolution Imagery, 18–20 November 2004, Pitsanulok, Thailand.
ESA, European Space Agency, 1995, Earth Observation Missions, ESRIN, Frascati.
ESA, European Space Agency, 1998, Envisat-1: Mission and System Summary, ESTEC. Hollandia Offset, Noordwijk.
Eurimage 1996, ERS, JERS-1 and Resurs-01 Missions, Training Courses on ERS SAR and Other Complementary Spaceborne Sensors. ESA-ESRIN, Frascati.
Filin S., Pfeifer N., 2006, Segmentation of airborne laser scanning data using a slope adaptive neighborhood. ISPRS Journal of Photogrammetry and Remote Sensing, 60(2): 71–80.
Flood M., Gutelius B., 1997, Commercial implication of topographic terrain mapping using scanning airborne laser radar. Photogrammetric Engineering and Remote Sensing, 63: 327.
Fowler R.A., 2000a, The Lowdown on LIDAR, Earth Observation Magazine, Marzo, p. 27.
Fowler, R.A., 2000b, LIDAR for Flood Mapping. Earth Observation Magazine, Luglio.
Fowler, R.A., 2000c, LIDAR Versus RADAR. Earth Observation Magazine, Settembre, p. 29.
Fritz, L.W., 1996, The era of commercial Earth observation satellite. Photogrammetric Engineering and Remote Sensing, 1: 39–45.
Fussel J., Rundquist D., Harrington J.A., 1986, On defining remote sensing. Photogrammetric Engineering and Remote Sensing, 52: 1507–1511.
Gianinetto M., 2006, Geocoding simulation of EROS-B synchronous imagery and comparison to EROS-A1 asynchronous data. Rivista Italiana di TELERILEVAMENTO – 2006, 36: 81–92.
Gomarasca M.A., 2000, Introduzione a Telerilevamento e GIS per la Gestione delle Risorse agricole e Ambientali, Ed. AIT, p. 250, 32 Tavole a colori; 2nd ed.
Hinz A., Dorstel C., Heier H., 2000, Digital modular camera: system concept and data processing workflow. Proceedings of ISPRS, Vol. XXXIII, Working Group II/7, Amsterdam.
Kyoto Protocol, 1997, Adopted at the Third Conference of the Parties to the United Nations Framework Convention on Climate Change (UNFCCC-COP 3) in Kyoto, Japan, 11 December 1997.
Kraus K., Pfeifer N., 1998, Determination of terrain models in wooded areas with airborne laser scanner data. International Journal of Photogrammetric and Remote Sensing, 53: 193–203.

International UAV Community, 2006. Referenced UAV Systems. UAV SYSTEMS: Global Perspective 2006, p. 171, www.uvs-info.com.

Marino C.M., 1993, Progetto LARA, Laboratorio Aereo per Ricerche Ambientali, Consiglio. Nazionale delle Ricerche, Pomezia, Roma.

Mastracci C., 1999, Introducing the Living Planet Programme, Atti 3° Conferenza Nazionale ASITA, Napoli 9-12 novembre, Vol. 1: LXXXV–LXXXIX.

NASA, National Aeronautics Space Administration, 1993, EOS References Handbook. Washington DC, USA.

NRSA, National Remote Sensing Agency, 1995, IRS-1C Data Users Handbook, Hyderabad, India, p. 1–169.

Shannon C.E., 1949, Communication in the presence of noise. Proceedings Institute of Radio Engineers, 37, no. 1, 10–21. Reprint as classic paper in Proceedings of the IEEE, 86, no. 2, Feb. 1998.

Obert H., 1923, Die Rakete zu den Planetenraumen (The Rocket into Planetary Space), p. 92.

Obert H., 1929, Die Rakete zu den Planetenraumen (The Rocket into Planetary Space), p. 429.

Pfeifer N., Lichti D., 2004, Terrestrial Laser Scanning Developments, Applications and Challenges, GIM International, 18(12): 50–53.

Pfeifer N., Kraus K., 2001, Advanced DTM generation from Lidar data. International Archives of Photogrammetry and Remote Sensing, XXXIV-3/W4: 23–35.

Rast M., Berger M., Sivestrin P., Del Bello U., 1999, Scientific Objectives and Derive Systems Requirements of the European Space Agency Earth Explorer, Land-Surface Processes and Interactions Mission, EUROPTO Conference, Florence, I, September, SPIE, Vol. 3870, pp. 49–57.

Rutzinger M., Höfle B., Pfeifer N., Geist T., Stötter H., 2006, Object-based analysis of airborne laser scanning data for natural hazard purposes using open source components. 1st International Conference on Object-Based Image Analysis, Salzburg, Austria.

Sapeta K., 2000, Have you seen the light: LIDAR technology is creating believers. GEOworld, October:26.

Schanda E., 1986, Physical Fundamentals of Remote Sensing. Springer-Verlag, Berlino, Germany.

Schulz T., Ingensand H., 2004, Terrestrial Laser Scanning – Investigations and Applications for High Precision Scanning, FIG Working Week, Athens, Greece, May 22–27, 2004.

Short N.M., 1982, The Landsat Tutorial Workbook, NASA Publication 1078. Government Printing Office, Washington DC, USA.

Seige P., 1995, MOMS-02/D2 mission. Proceedings of MOMS-02 Symposium, Cologne, Germany, 5–7 July, pp. 41–51.

Sohn H.G., Yoo, H.H., Kim, S.S., 2004. Evaluation of geometric modelling for KOMPSAT-1 EOC imagery using ephemeris data. ETRI Journal, 26(3): 218–228.

Thibault D.A., 1995, Land satellite information in the future. Proceedings Land Satellite Information in the Next Decade, Sett. 95, Vienna, Virginia, USA, III: pp. 14–21.

Tomasi C., Prodi F., Sentimenti M., Cesar G., 1983, Multiwavelenght sunphotometers for accurate measurements of atmospheric extinction in the visible and near-IR spectral range. Applied Optics, 22: 622–630.

Turner A.K., 2000, LIDAR provides better DEM data. GEOworld, November: 30.

Tyler W.A., 1993, The Multispectral Scanner, Landsat Other Sensor, Earth Observing Magazine.

Ulaby F.T., Moore R.M., Fung A.F., 1981, Microwave Remote Sensing: Active and Passive, Vol. I. Microwave Remote Sensing, Fundamentals and Radiometry. Addison-Wesley Publishing Company, Reading, Massachussetts, USA.

Ulaby F.T., Moore R.M., Fung A.F., 1982, Microwave Remote Sensing: Active and Passive, Vol. II, Radar Remote Sensing and Surface Scattering and Emission Theory. Addison-Wesley Publishing Company, Reading, Massachussetts, USA.

Ulaby F.T., Moore R.M., Fung A.F., 1986, Microwave Remote Sensing: Active and Passive, Vol. III, From Theory to Applications. Addison-Wesley Publishing Company, Reading, Massachussetts, USA.

Volz F.E., 1959, Photometer mit Selen-Photoelement zur spektralen Messung der Sonnenstrahlung und zur Bestimmung der Wellelangenabhangingkeit der Dunsttrubung. Archives for Meteorology, Geophysics and Bioclimatalogy., 10: 100–131.
WegmŸller U., 1995, Land surface analysis using ERS-01 SAR. ESA Bulletin, 81: 30–37.
WegmŸller U., Werner C.L., NŸesch D., Borgeaud M., 1997, Retrieval of vegetation parameters with SAR interferometry. IEEE Transaction on Geoscience and Remote Sensing, 35, 1: 18–24.
Wehr A., Lohr U., 1999, Airborne laser scanning—an introduction and overview. ISPRS Journal of Photogrammetry and Remote Sensing, 54(2–3): 68–82.

Chapter 7
Satellite Positioning Systems

This chapter describes the main satellite positioning systems available worldwide. More prominence is given to the NAVSTAR GPS: *NAVigation Satellite Timing and Ranging Global Positioning System*, a system developed in the United States and adopted by most countries in the world.

Successively a brief description of the GLONASS, *Global'naya Navigatsionnaya Sputnikovaya System*, developed by the former Soviet Union and now managed by Russia, is presented.

Galileo is the most recent experience in the GPS field. It is an ongoing programme developed by the European Community aimed at equipping Europe with its own global positioning system needed to reduce the dependence on the GPS system for strategic and economic reasons. *Galileo* opens the way for new positioning services and is both alternative and complementary to the American and Russian systems. The European Council adopted on July 19, 1999 a solution which opened the way to the *Galileo* system in cooperation with the European Space Agency (ESA).

7.1 NAVSTAR Global Positioning System (GPS)

The American satellite global positioning system is the NAVSTAR GPS. Its birth, supported by the American Department of Defence, goes back to 1973, when it became operative as a system devoted to serving military ships during navigation: it provides real-time position according to the WGS84 (*World Global System 1984*) geodetic reference system, valid for the whole Earth. Civilian utilization mainly related to survey and navigation application occurred some years later, since the early 1980s.

The system can be thought as comprising three modules, or segments:

- the space segment;
- the control segment;
- the survey segment.

The *space segment* is represented by the satellite constellation. This, in the case of NAVSTAR GPS, is made up of 24 operative satellites and some reserve ones, orbiting at an altitude of 20,183 km with a period of 11 h and 56 min. The constellation, completed by the consequent launches of several blocks of satellites, guaran-

M.A. Gomarasca, *Basics of Geomatics*, DOI 10.1007/978-1-4020-9014-1_7,

tees the simultaneous visibility of at least four satellites (number strictly necessary for the 3D positioning), at any time in any place on the Earth. All the visible satellites present an elevation higher than 15°.

Satellites orbit on six different orbital planes that are 55° inclined with respect to the equatorial plane and 60° rotated with respect to the other (Plate 7.1). Due to the difference between *sidereal time*, referred to the satellite, and *solar time*, referring to the observation point on the Earth, the satellites appear in the same point 4 min in advance every day, which determines a 2 h advance in the coverage every 30 days. Each satellite is equipped with four atomic clocks, two caesium and two rubidium, able to control the frequencies and the modulations for a steady and definite time reference equal to 10^{-12} to 10^{-14} s. Future blocks of satellites will be equipped with MASER-type clocks having a higher stability equal to 10^{-14}–10^{-15} s.

The constellation is programmed to be considerably improved thanks to 18 new satellites of the IIF block: a higher number of satellites will allow a better coverage and consequently a better performance in the signal reception; the improved features of the next-generation satellites will permit them to manage any frequency variation and to exploit measurements coming from the on board inertial systems (Fig. 7.1).

This segment is in charge of transmitting the signal to the ground, to maintain the time reference, to receive and execute the instructions coming from the control segment.

The *control segment* is represented by five ground stations (the main one is based in Colorado Spring, United States) distributed along the equator. It is in charge of:

- tracking the satellites through the computation and monitoring of the ephemeredes;
- controlling the on board clocks (through a hydrogen MASER clock);
- correcting, if needed, orbits;

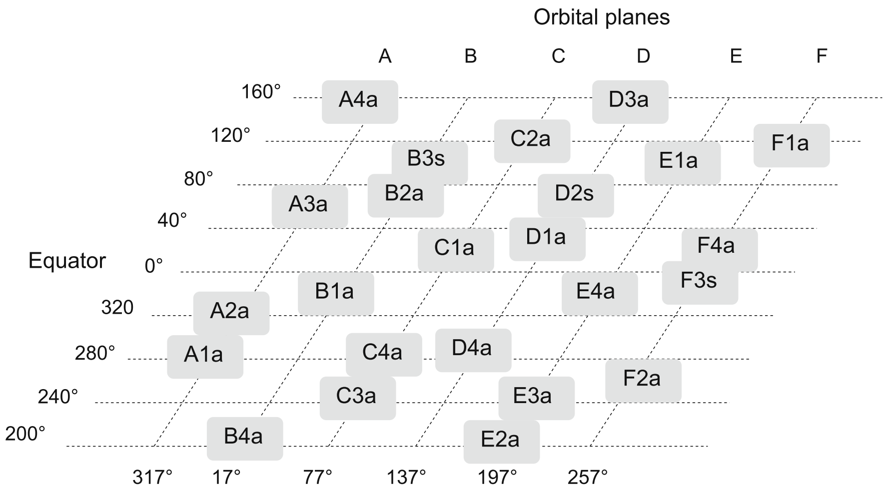


Fig. 7.1 Orbital distribution of the 24 NAVSTAR satellite. S: reserve satellite; (**a**) active orbital position

- storing new data on the satellites (foreseen ephemeredes) useful for users and transferred to the ground receivers through the GPS signal (navigation message).

The *Ephemeredes* are defined as numeric tables providing the coordinates of astronomical bodies such as the Sun, the Moon, the Planets and artificial satellites, which change their position during a given time period (Box 7.1).

The *survey segment* is represented by a device that receives the transmitted signal through an antenna; the receiver must be placed at the point whose coordinates are intended to be determined (Fig. 7.2).

Box 7.1 Keplerian parameters to define the coordinates of the satellite used to compute the ephemeredes

Keplero parameter	Symbol	Specification
Orbit inclination	i	Angle between the orbital plane and the equatorial plane
Longitude of ascending node	Ω	Orbital elements used to specify the orbit of an object in the space. For a Sun-orbiting body, it is the angle formed at the Sun from the First Point of Aries to the body's ascending node, measured in the reference plane (the ecliptic) and in the direct sense
Longitude of the perigee (argument of periapsis or of perihelion)	ω	Angle between the ascending node (the point where the orbiting body passes from the southern to the northern hemisphere) and the periapsis (the point of the closest approach to the central body), measured in the body's orbital plane and in its direction of motion. It is undefined for equatorial orbits, where there is no defined ascending node, and for circular orbits, where there is no defined periapsis
Orbit eccentricity circular orbits $e = 0$ elliptic orbits $0 < e < 1$ parabolic trajectories $e = 1$ hyperbolic trajectories $e > 1$	e	Important parameter of the orbit that defines its absolute shape. Eccentricity can be interpreted as a measure of how much this shape deviates from a circle
Semi-major axis of the orbit	a	The major axis of an ellipse is its longest diameter. The semi-major axis is one half of the major axis and thus runs from the centre, through a focus, and to the edge of the ellipse
Orbital period (mean anomaly at epoch)	P_b (M_o)	Time to make one full orbit In the study of orbital dynamics, the mean anomaly is a measure of time, specific to the orbiting body *p*, which is a multiple of 2π radians only at periapsis. It is the fraction of the orbital period that has elapsed since the last passage at periapsis *z*, expressed as an angle
Julian date of the perigee (passage at the perigee)	T_o	The perigee is the nearest point to the Earth

Fig. 7.2 World geographical dislocation of the control stations: (**a**) Hawaii; (**b**) Colorado Springs; (**c**) Ascension; (**d**) Diego Garcia; (**e**) Kwaialein

7.1.1 The GPS Signal (NAVSTAR)

The data coming down to the ground are transmitted by the satellites as a complex coded signal modulated by the L_1 and L_2 frequencies: they are determined as multipliers of the fundamental frequency f_0, equal to 10.23 MHz, produced by the on board clocks. Modulation operates over the transmitted code. The P-code (more precise) is modulated on both the frequencies while the C/A code (less precise) is modulated only on the L_1 frequency (Table 7.1).

The L-band *Standard Positioning Service* (SPS) ranging signal is contained within a 2.046 MHz band centred on L_1. The carrier frequency for L_1 is coherently derived from a frequency source inside the satellite. The nominal frequency of this source, as it appears to an observer on the ground, is 1.023 MHz. To compensate for relativistic effects, the output frequency of the satellite frequency standard, as it would appear to an observer located at the satellite, is 10.23 MHz offset by:

Table 7.1 GPS signal frequencies

Frequency type	Frequency (MHz)	Wavelength (λ)
Fundamental (satellite's frequency standard)	L : 10.23	
Carrier frequency	L_1: 1575.42 $L_1 = L \times 154$	λL_1 : 19.0 cm
	L_2: 1227.60 $L_2 = L \times 120$	λL : 24.4 cm
Impulsive	P : 10.23	λP : 30 m
	C/A : 1.023	$\lambda C/A$: 300 m
Navigation message	D : 50 Hz	

$$\mathrm{Df/f} = -4.4647 \times 10^{-18} \text{ or}$$
$$\mathrm{Df} = -4.567 \times 10^{-3}\,\mathrm{Hz}$$

This frequency offset results in an output of 10.22999999543 MHz, which is frequency divided to obtain the appropriate carrier modulation signal (1.022999999543 MHz). The same output frequency source is also used to generate the nominal L_1 carrier frequency (f_o) of 1575.42 MHz.

The L_1 carrier frequency, 1575.42 MHz, contains the C/A-Code, the encrypted P-Code (or Y-Code) and the *navigation message*. Commercial GPS navigation receivers can track only the L_1 carrier to make pseudo-range (and sometime carrier phase and Doppler frequency) measurements.

The L_2 carrier frequency, 1227.60 MHz, contains the encrypted P-Code (or Y-Code) and the *navigation message*. Military Y-Code capable receivers can, in addition to L_1 measurements, make pseudo-range measurements on the L_2 carrier. The combination of the two measurements (on L_1 and L_2) permits the correction of the *ionospheric delay*. Dual-frequency GPS receivers, intended for surveying applications, can make L_2 measurements using proprietary signal processing techniques. Such measurements are essential if the *ionospheric delay* on carrier phase is to be corrected (especially on baselines of length greater than 20–30 km) and/or where fast ambiguity resolution is needed.

Both the codes (P, C/A), needed to measure the satellite-to-receiver distance, consist in a pseudo-casual sequence of +1 and –1 which is repeated after a predefined interval (equal to 1/1000 s for the C/A code and 1 week for the P-code). An information code D is modulated on both the frequencies: it contains the *navigation message* with reference to parameters of correction of intervals and ionospheric effects, forecasted ephemeredes, almanac and satellite operating state, and other codes and a range of alpha-numerical messages.

7.1.2 GPS Measurement

The basic measurement performed by the GPS technology is the distance between the satellite and the receiver. This is derived indirectly through auxiliary measurements relating to both the signal transmission from the satellites and the receiver recording dynamics (Fig. 7.3).

A copy of the received signal is obtained inside the receiver, through the oscillator. The received and the generated signals are then compared to compute their displacement related to the signal period from the satellite to the receiver (antenna). After that the GPS signal can be treated according to two different strategies:

- *pseudo-distance* or *code* measurement, allowing the real-time positioning;
- *phase* or *displacement* measurement, carried on the frequencies L_1 and L_2 that requires post-processing of the data.

In both the situations, the critical point is the signal period (Fig. 7.4). The accuracy of the time measurement is the conditioning factor of the positioning quality. Once the signal period is known, the satellite-to-receiver distance can be determined according to the following simple relation:

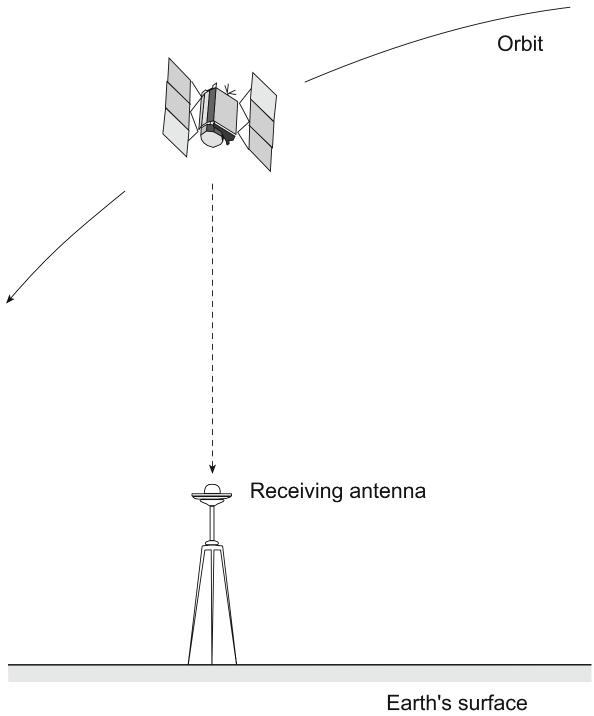


Fig. 7.3 Measurement of the receiving station–satellite distance

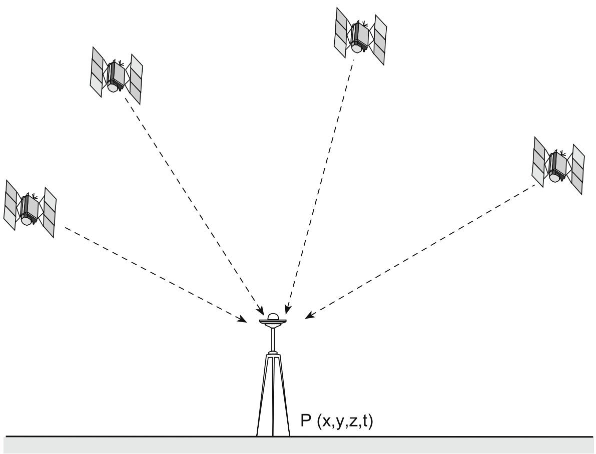


Fig. 7.4 Global positioning systems: the code and phase measurements received from at least four satellites are used to compute the spatial coordinates (X, Y, Z). The fourth satellite solves the temporal incognita (t) of the system

$$\Delta S = t \cdot c$$

where

ΔS : satellite-to-receiver distance
t : signal period (~ 0.070 s)
c : speed of light in the vacuum 2.99×10^{10} ($cm \cdot s^{-1}$)

Satellite and receiver clocks are not synchronized. Moreover, they suffer different synchronism errors compared to a fundamental reference time scale.

The synchronism error, referred to the receiver fundamental scales, is about 1 ms. Considering the signal propagation speed (2.99×10^{10} $cm \cdot s^{-1}$), such uncertainty determines a potential error in the satellite-to-receiver distance measurement of about 300 km.

The synchronism error is about 1000 times smaller (˜1 ns) than the satellite clock fundamental scale; this produces a measurement error of about 30 cm, reasonably negligible. The satellites carry atomic clocks (caesium and rubidium), while the receivers host quartz clocks.

7.1.2.1 Code or Pseudo-distance Measurements

Pseudo-distance measurements relate to the signal period from the satellite to the receiving ground station; the signal propagation (speed of light) is known, thus the satellite-to-receiver distance can be calculated. The four distances measured at the same time by four satellites can be combined through a multiple spatial intersection to calculate the 3D position of the surveyed ground point. The measurement accuracy, as obtainable after May 2, 2000, depends on the type of code used:

- the C/A code-based calculation suffers an uncertainty of ± 5–10 m;
- the P-code based calculation suffers an uncertainty of ± 1–3 m.

These methods are based on the comparison between the signal received from the satellite and the signal generated by the receiver. The code inside the signal is defined as an appropriate sequence of bits (one and zero), such that it looks almost casual, similar to a noise better than a real signal. For this reason, it is defined as *pseudorandom code*, or *pseudo-range*, or *pseudo-distance*. The comparison between the two signals defines a displacement on the bit sequence that can be used to measure the satellite signal period (or delay): the above-mentioned synchronism errors represent a further unknown quantity to be determined (Fig. 7.5).

7.1.2.2 Phase Measurement

In the early 1980s, two radio-astronomers, Counselman and Shapiro from the Massachusetts Institute of Technology, tested an alternative at the *pseudo-distance*

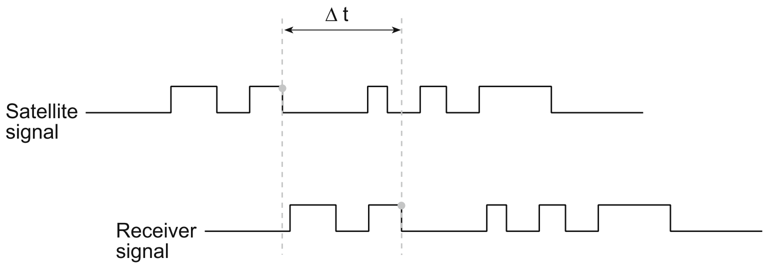


Fig. 7.5 Time difference (Δt) between the signal from the satellite and the signal generated by the receiver on the bit sequence. The two signals are identical but with a phase difference

(pseudo-range) modulation, using the same principle adopted in the radar interferometric measurements (see Chapter 4): *phase measurements* instead of *displacement measurements* (carrier frequency). The phase measurements, with calculation of the distance between antenna and object, have been successfully used for years in the diastimeters and with the radio signals for extra-galactic tests. This approach enables an important increase of the positioning accuracy.

In phase measurements, the observation is between the phase of the demodulated carrier wave (sinusoidal L_1 and sometime L_2) from the satellite and the phase of the reference wave generated by the receiver at the same time. In calculating the distance, it is important to quantify:

- the fractional part of the displacement that can be derived by a phase discriminator with a precision of the order of 1% of the signal wavelength;
- the number of entire cycles elapsed (*phase ambiguity*), whose computation is possible during the data post-processing. The satellite-to-receiver distance is then obtained as a multiple (entire part + fraction) of the wavelength of the considered signal.

7.1.3 GPS Operative Mode

The GPS system, as noted above, can be successfully used for geodetic–topographic surveys and for maritime, aerial or terrestrial navigation. According to the application, GPS data must be acquired and processed in different ways. Possible survey strategies are

- *absolute positioning*: point coordinates are determined in a global reference system;
- *relative positioning*: the components of the vector connecting two points are determined. This way the systematic errors in the two ground stations can be eliminated;
- *differential positioning*: satellite-to-receiver measurements are corrected on the basis of a reference station whose coordinates are known (see Fig. 7.11).

In the first case, the absolute coordinates of a single point at a time can be determined, while in the relative and differential positioning coordinates of a point are defined relative to another point selected as reference.

For each positioning strategies, different measurement modes, related to their duration, can be considered:

- *static*: if the receiver remains on a point for a long period;
- *kinematical*: if the receiver continuously moves (navigation).

In addition, measurement can be done:

- *in a code mode*: if the processing refers to the signal pulse component;
- *in a phase mode*: if the processing refers to the signal carrier.

The survey can be an opportune combination of these three groups of possibilities (i.e. absolute positioning with code measurements, either static or kinematics phase measurements):

- *absolute with code measures*: the signal pulse component C/A or P is used. The interval between the transmission of the signal from the satellite to the receiver is measured;
- *absolute with phase measures*: the measurement of the satellite-to-receiver distance is performed using the sinusoidal carriers frequencies L_1 and L_2 in different modes (not in use);
- *relative static*: it is necessary to simultaneously operate with at least two permanent receivers for the period of the survey. The purpose is to measure the components of the vector between the two points A–B, or the baseline AB: a receiver must be positioned simultaneously on each point. For this purpose, both code and phase measurements can be used; the latter are the most precise even if the least practical.
- *kinematics relative*: it is obtained by placing a receiver in a point (*base*) and moving the other receiver (*rover*) along the path of interest acquiring data continuously;
- *differential GPS (DGPS)*: at least two receivers are used: one is permanent (*base*) in a station with known coordinates; a second (*rover*) is positioned on the point to be detected. The second receiver can be moving. The DGPS can be applied to both code and phase measurements. This measurement improves the precision of the coordinates of the point by means of correction derived from the precise coordinates of the base. The difference between the satellite-to-receiver distance (*range*) and the distance calculated on the base station is assessed; the correction is applied to the distance between each satellite and receiver. The positioning of the *rover* is then determined by using the appropriate satellite-to-receiver distance (Fig. 7.6).

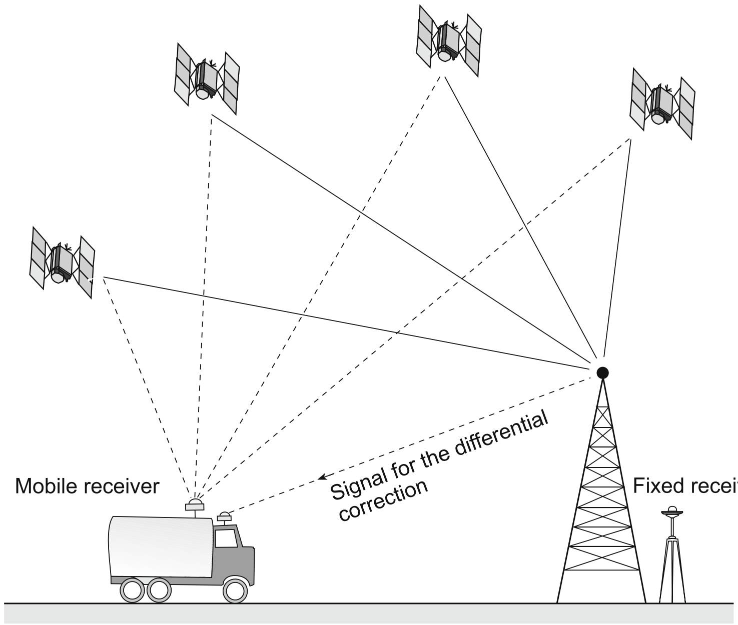


Fig. 7.6 Real-time differential GPS

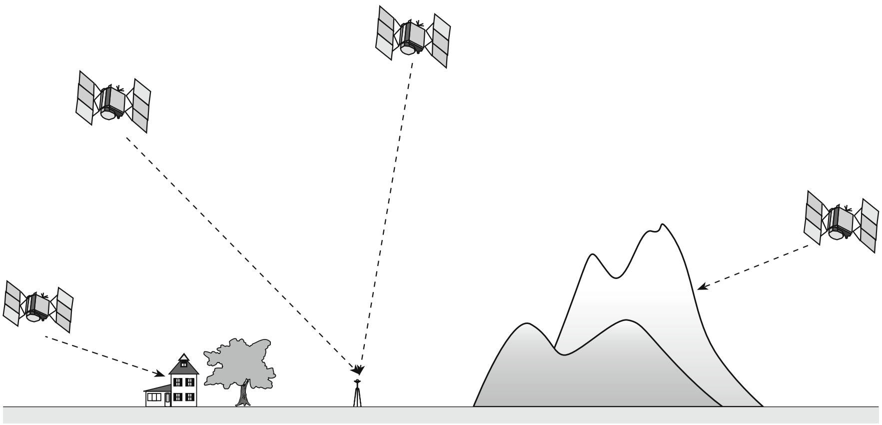


Fig. 7.7 The positioning systems operate if at least four satellites are visible by the receiver–antenna system. The signal can be stopped (or interfered with) by the presence of buildings, trees and mountains

Though GPS can really modernize the surveying of ground points for geodetic purposes, some limitations exist especially related to the presence of a great variety of obstacles that can strongly affect their functioning: the signal is, for example, completely absent or irregular in urban areas, in dense woods and in mountain areas (Fig. 7.7).

7.1.4 GPS Errors

The GPS system is affected by different kinds of errors which can be completely or partially removed adopting some precautions and technical solutions. The errors may be present in the measurements carried out by the operator: they can be systematic or induced by the system management, as for the Selective Availability (SA).

7.1.4.1 Selective Availability (SA)

The selective availability is a policy adopted by the *United States Department of Defence* to introduce some intentional clock noise into the GPS satellite signals thereby degrading their accuracy for civilian users. This produces a circular error up to a 100 m making the ground measurements inaccurate. To overcome this limitation, the differential GPS (DGPS) was developed and a local ground station added as support at the satellite-based stations, improving the accuracy. Selective availability contains GPS clock dithering that heavily impacts on the baseline solution; the arbitrary acceptance of the broadcast ephemeredes is ill-advised.

Starting from 04:05 UTC (*Universal Time Coordinated*) on May 2, 2000, SA has been removed by the US government with a consequent measurement improvement (up to 10 factor). GPS receivers for civilian use, which use the C/A (*Coarse Acquisition*) low-resolution signal, before SA suppression provided position within a range of about 100 m. After suppression the potential error ranged from 5 to 10 m according to the recording conditions (Fig. 7.8).

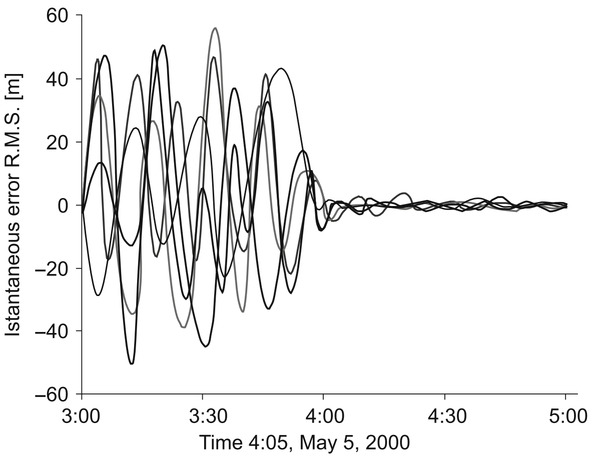


Fig.e 7.8 The signal registered from a GPS receiver before and after the removal of the selective availability (SA). The increased resolution improved the GPS autonomous positioning of a10 factor (van der Marel, 2000)

Table 7.2 Altimetric and planimetric positioning errors. Signal recorded by a GPS station in The Netherlands at the UTC reference time. The clearance of the Selective Availability (SA) improved the GPS autonomous positioning of a 10 factor

May 2, 2000	R.M.S. error (m)		95 % Percentile (m)	
	Horizontal	Vertical	Horizontal	Vertical
0:00–4:05	32.0	56.0	66.0	109.0
4:05–8:10	3.1	5.4	5.2	12.0

The signal liberalization has reduced the use of the DGPS for a wide range of users, as its main function was to remove the errors due to the selective availability, as well as to reduce some errors due to the system (Table 7.2). For applications requiring a tolerance more than or equal to 10 m for planimetry and 20 m for altimetry, the *pseudo-distance* or *code* measurements, with a single instrument and without any differential correction, can be appropriate (Figs. 7.9 and 7.10).

The new situation makes less important the use of satellite signals from the Russian system GLONASS (*Global Navigation Satellite System*) which, being free from induced errors like SA, in the past permitted the obtaining of a precision of about 10 m. The synergy with GLONASS satellites is nevertheless still useful because the higher the number of available satellites for the intersection of the distances, the higher the possibility of obtaining the position of points in urban or hilly areas.

The removal of the selective availability is part of a strategic programme of gradual replacement of the satellites making up the constellation. The new ones will also have a third functioning frequency and new access modes for military use with a higher precision and reliability.

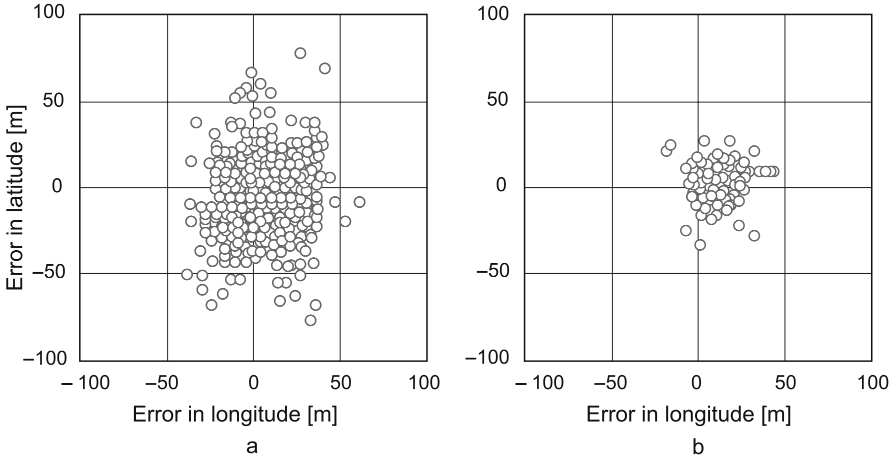


Fig. 7.9 Planimetric (2D) positioning errors before (**a**) and after (**b**) the liberalization of the signal. ΔN: planimetric positioning error in north coordinate; ΔE: planimetric positioning error in east coordinate

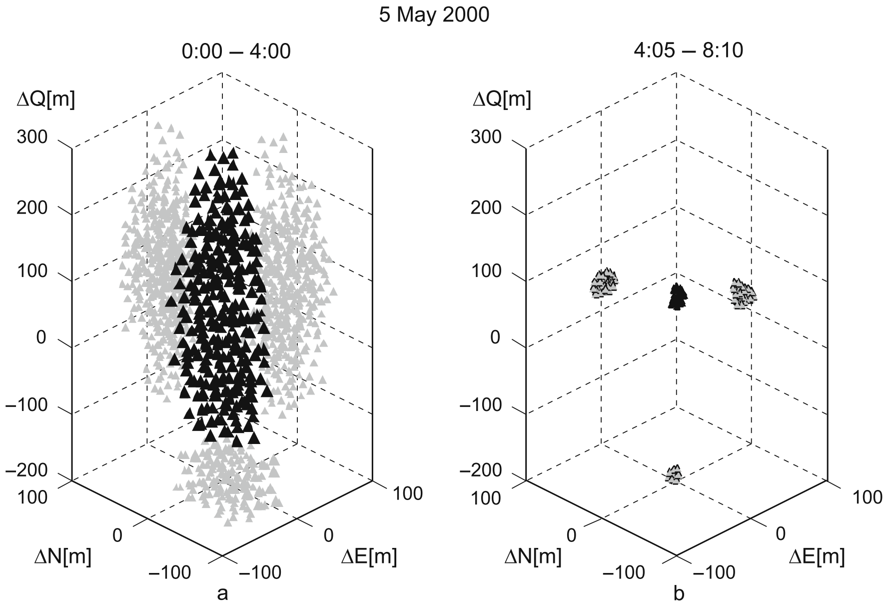


Fig. 7.10 Three-dimensional (3D) positioning errors before (**a**) and after (**b**) the liberalization of the signal. ΔQ: altimetric positioning error; ΔN: planimetric positioning error in north coordinate; ΔE: planimetric positioning error in east coordinate

7.1.4.2 Systematic Errors

The positioning accuracy which is obtained by GPS measurements increases along with the number of satellites simultaneously dialoguing with the receiver; thus it is important that the antenna is positioned as far as possible in open areas, i.e. free from obstacles for elevation angles higher than 15–20° (necessary to reduce tropospheric delays).

Errors affecting GPS measurements can be accidental or systematic.

Satellite errors are

- clock errors;
- errors in calculation of satellites position;
- signal propagation through the atmosphere (tropospheric and ionospheric errors).

Ground instruments errors are

- antennas;
- receivers;
- environmental conditions.

7.1.4.3 Accidental Errors

The theoretical precision achievable in the positioning (Table 7.3) can be estimated equal to 1/100 (1%) of the wavelength (λ).

Table 7.3 Theoretical precision in the positioning

GPS signal	λ (m)	Precision (m)
Carrier frequency L_1	0.19	±0.002
Carrier frequency L_2	0.244	±0.002
Code P	30	±0.3
Code C/A	300	±3

7.1.4.4 Orbital Errors

At the basis of an absolute positioning of a point by GPS is the knowledge of the satellite ephemeredes, which are defined with precision of about 50 m if forecast and about 1 m if measured a posteriori. These errors are less relevant if the positioning is differential or relative.

7.1.4.5 Troposphere and Ionosphere Errors

The troposphere, i.e. the part of the atmosphere up to 40 km, and the ionosphere, above the troposphere from 40 to 1000 km, generate a delay in the signal propagation increasing the measure regarding the satellite-to-receiver distance (see Fig. 4.16).

The tropospheric delay is independent of the frequency and thus it is identical both for the L_1 and L_2 carrier phases. On the contrary, it depends on the atmospheric parameters and the satellite zenith angle (z). When the satellite is seen with an elevation lower than 15° with respect to the horizon, corresponding at 75° zenith angle, the distance error due to the refraction effect unacceptably increases. Hence it is necessary to consider, during the intersection of the distances, those satellites having an elevation of at least 15°.

The ionosphere induces errors depending on the frequency of the signal. Errors are, therefore, different for L_1 and L_2 carriers. An appropriate combination of the two carrier phases almost entirely removes the effect due to the ionospheric delay in the case of measurements having base longer than 15 km; with shorter base ionospheric effect can be removed by differential or relative methods.

Systematic error effects can be reduced or removed through relative methods, or *equation differentiation* techniques, in which the differences of phase measurements are considered. These techniques often induce increase of noise of measurements.

Single frequency receivers generate errors due at Sun-induced phenomena in the ionosphere, not continuous in time. Moreover, they have a daily cycle with a peak about 2 h after midday, local time; the error due to the ionosphere is higher near the equator. At medium latitudes, like in Italy, this error can be estimated as 20–50 m at 12.00 am, according to satellite elevation: higher orbits generate smaller errors. The models used for the correction of the ionospheric errors can partially (about 50–60%) solve the problem. Only simultaneous consideration of the frequencies L_1 and L_2 may further reduce the effects of the ionosphere on the positioning accuracy. The latter depends both on the precision of the measurements and the geometric distribution of the satellites with respect to the ground point. Hence a *Dilution of*

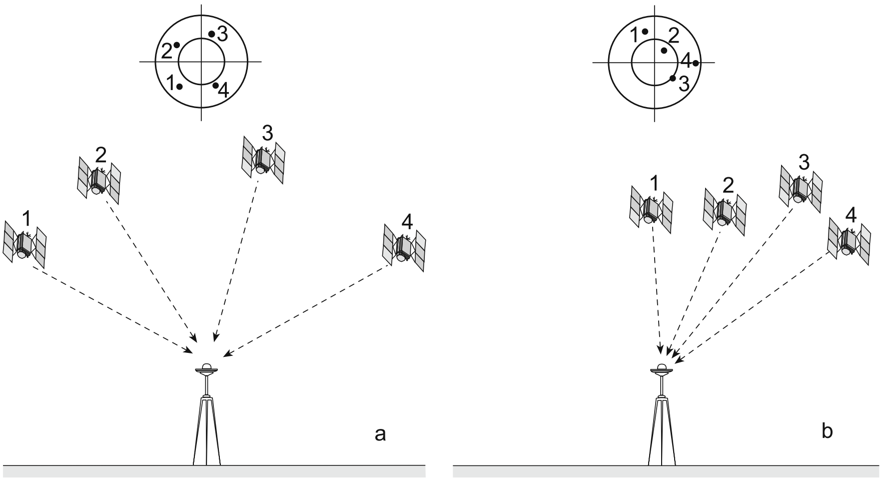


Fig. 7.11 The GDOP, i.e. 3D positioning accuracy, related to the constellation geometry; (**a**) is lower than in the constellation (**b**), both the constellations being constituted by four satellites. The volume of the resulting pyramid, (**a**) whose base's vertexes are the four satellites, is bigger than the pyramid in (**b**)

Precision (DOP) index was introduced; it represents the contribution of the satellite geometric distribution to the positioning accuracy. There are different types of DOP according to the considered coordinates (PDOP, GDOP, HDOP, VDOP and TDOP). In particular the GDOP, index measuring the geometric suitability of the satellite configuration, is inversely proportional to the volume of a solid having as vertices the satellites and the ground point (Fig. 7.11).

7.1.5 GPS Geodetic Reference System

GPS operates in a Cartesian reference system called WGS84, defined as follows:

- system origin is the Earth's centre of mass (geocentric system);
- z-axis is parallel to the direction of the conventional terrestrial pole;
- x-axis defined as the intersection of the equatorial plane (orthogonal to the z-axis) with the plane of the Greenwich meridian (0° meridian);
- y-axis defined as orthogonal to the x- and z-axes, to complete the clockwise Cartesian system.

The ellipsoid associated at this system, having its centre and rotation axis respectively coinciding with the origin and the z-axis of the Cartesian system, can acceptably approximate the geoid (average differences ± 50 m). The coordinates obtained from GPS observations may be expressed both in Cartesian coordinates (x, y, z) and in ellipsoid (geographic) coordinates φ, λ, H (latitude, longitude and altitude with respect to the ellipsoid), or in plane cartographic coordinates.

As the reference system adopted for geodetic, topographic and photogrammetric surveys is usually a hybrid system referring to national data for planimetry and to the geoid for altimetry, some coordinate transformations are needed. As far as planimetry is concerned, the transformation can be defined through the definition of the parameters of a 3D scaled roto-translation (three for translation, three for rotation and one for scale). The transformation parameters are local, and they can be estimated according to some points whose coordinates are known in both the systems (at least two points for planimetric transformations, at least three for plano-altimetric transformations). Concerning altimetry, the GPS data refer to an ellipsoidic surface while the required altitude is usually referred to the geoid. The distance separating the two reference surfaces ellipsoid and geoid is called *Geoid Undulation*(N); it is calculated at global scale and refined at local level, in order to better model local anomalies and topography (see Fig. 2.12).

It is worth emphasizing that the reference system is unusable if the coordinates and the stability of the points of a persisting network are not defined and materially signalized (*International Earth Reference Frame*, ITRF). The *European Reference Frame* (EUREF) considers a continental average deformation speed, useful to define a local reference system. The reference system adopted by EUREF is the same of the ITRS (*International Terrestrial Rotation Service*) as calculated in 1989. It is fixed to the steady part of the Euro-Asia tectonic plate and it is called *European Reference System 89* (ETRS89) (Plate 7.2).

7.1.6 Receivers

There are different categories of receivers whose characteristics depend on the frequencies they can receive and hence on the measures they can perform (Table 7.4):

- *code measure receivers* (pseudo-distance): they can acquire the C/A code, which is the pulsing component. They are the simplest and mainly used for sea and air navigation. They allow absolute positioning with accuracy of 10–30 m. Position coordinates are visualized by a small liquid crystal screen and they can also be stored;
- *single-frequency receivers*: they can acquire both the C/A code and the carrier L_1, enabling either code or phase measures. They permit static or cinematic positioning for limited basis (not longer than 15–20 km). Measurements can be stored and successively downloaded for post-processing;
- *double-frequency receivers*: they can receive the available frequencies of both the modulating (C/A and P) and the carriers (L_1 and L_2) components. They permit processing of the carriers to perform relative and differential positioning, both static and dynamic, without limit for the length of the base. They produce measurement with high precision.

Receivers can support other sensors for many different uses: inertial systems, laser scanners, photogrammetric cameras, geophysical instrumentation, precision farming and the traditional measurement of geodetic networks.

Table 7.4 Types of receivers of the GPS signal

Receiver	Frequency	Application	Precision
Code measurement	C/A	Navigation	10–30 m
Single frequency	C/A e L1	All	High
Double frequency	C/A P L_1 L_2	All	High

7.2 GLONASS Global Positioning System

There is another Positioning System designed by the former Soviet Union and managed by Russia: the GLONASS (*Global'naya Navigatsionnaya Sputnikovaya System*). It was programmed in the mid-1970s, and it has been operative since 1982. It is similar to the NAVSTAR GPS both in system design and in operational mode.

This satellite radio-navigation system enables an unlimited number of users to determine their own coordinates, their speed components and the synchronization with respect to the system time, everywhere on the terrestrial surface and in any atmospheric condition. The system is based on a constellation of satellites which continuously transmit signals codified on two frequency bands.

7.2.1 GLONASS Characteristics

The complete constellation of the GLONASS system is constituted by 24 satellites distributed on three orbital planes that are 120° equidistant from one another, each containing eight satellites, uniformly distributed with a step of 45°.

The orbit mean altitude is 19,100 km, with 25,510 km maximum (orbital radius). The rotation speed is near 14,000 km/h. The orbital planes inclination is 64.8°. The maximum shifting from its ideal position is not higher than ± 5° over a period of 5 years. The constellation is continuously monitored to check the satellites' position and their functioning status, collection of the ephemeredes data and determination of the displacement between the satellites and the GLONASS time.

The system time is given by hydrogen oscillators whose unstableness is lower than $5 \times \cdot 10^{-14}$ s. GLONASS time goes along with the UTC (*Universal Time Coordinated*), determining a discontinuity and a gap with the GPS system time scale which, instead, refers to the UT1 system (*Universal Time, corrected universal time*). The problem is solved every 6 months by appropriate corrections established by the *Bureau International de l'Heure* (BIH) that is in charge of supplying results to the users. The GLONASS system transmits on two main frequencies (1602 and 1246 MHz) of the L-band, characterized by a k factor for each one of the constellation 24 satellites:

$$L_1{:}1602 + k \cdot 0.5625\,\text{MHz}$$

$$L_2{:}1246 + k \cdot 0.4375\,\text{MHz}$$

where

k: variable ranging from 0 to 24 in correspondence with the broadcast channel typical of each satellite.

After some problems of proximity with the frequencies used by radio-astronomy for the observation of interstellar space, in order to avoid any interference, the number of dedicated frequencies was reduced to 12, plus 2 control frequencies, with respect to the previous 24+1. Frequencies were limited to the range 1598.06–1604.25 MHz for L_1 and to the range 1242.94–1247.75 MHz for L_2. According to this new configuration, the satellites on the same orbital planes separated by 180° latitude use the same channel without any interference effect. The relocation towards the new assignation began in 1993.

Navigation signals are transmitted according to two modes:

- SPS (*Standard Position Service*): it provides in real time the acquisition accuracy corresponding to 20–30 m in planimetry and 20–40 m in height, without any intentional degradation. Accuracy is lower than that derivable from the GPS system free from any strategic error or selective availability (SA). This code is transmitted by both the carrier L_1 and L_2.
- PPS (*Precise Position Service*): precision code transmitted only by the L_2 band. To access the PPS code, the authorization of the Russian Ministry of Defence is necessary.

The navigation message is updated every 30 min. The main transmitted parameters concern:

- satellite ephemeredes calculated from the ground control centre;
- on board clocks synchronisms with Moscow time;
- satellite position expressed with respect to the three Cartesian geocentric coordinates X, Y, Z;
- monitoring of the satellite speed;
- difference between the GPS time and the GLONASS time, in connection with the GPS navigation message, to allow synchronization and corrections between the two systems if used together.

GLONASS satellites ephemeredes are very simple as they are already expressed in terms of position, speed and Cartesian acceleration, while in the GPS system the orbital parameters must be transformed into Cartesian geocentric coordinates.

For a correct intersection, it is necessary that the receiver can see at least four satellites simultaneously. The receiver to determine the position must solve a four-equation linear system in four unknowns: the geocentric coordinates X, Y, Z and the time difference between the receiver clock and the GLONASS system. The computation is carried out considering the main errors, even including ionospheric and tropospheric effects. The Cartesian geocentric coordinates are referred to the *Parametry Zemli 1990* local system (PZ90). The parameters of this local reference system are shown in Box 7.2.

Box 7.2 Parameters of the local reference system PZ90 referred to the geocentric Cartesian coordinates of the GLONASS system

Parameter	Unit	Value
Velocity of the terrestrial rotation	radians/s	7.292115×10^{-5}
Gravitational constant	m^3/s^2	398,644.44109
Atmosphere gravitational constant	m^3/s^2	0.35109
Speed of light in vacuum	m/s	299,792,458
Semi-major axis	m	6,378,136
Compression	ratio	1/298.257
Equatorial gravity acceleration	mgal	978,032.8
Atmospheric correction of the gravity acceleration at the sea level	mgal	–0.9

Three ways can be followed to transform the GLONASS coordinates system (PZ90) into the GPS system (WGS84):

- comparison between the absolute coordinates of ground stations expressed in both the systems;
- comparison of the satellite ephemeredes expressed in both the coordinates system;
- comparison of the relative positioning results of both the coordinates systems, in case of a four-parameter transformation: three angles and one scale factor.

7.2.2 GLONASS Versus NAVSTAR GPS

GLONASS has the advantage of having up to 14 satellites visible above the horizon at the same time, while GPS has only 7–8: this feature allows a higher precision in complex environmental conditions like hilly areas and valleys or city centres.

GLONASS never had, intentionally, disturbed signals due to military or strategic (SA) reasons like the GPS. This fact always guaranteed better positioning opportunities, comparable with the current GPS, without any differential correction, to have just a few metre error positioning using a single receiver.

Moreover GLONASS offers a wider and better covering for the highest latitudes, thanks to a different disposition in the satellites' configuration: orbits are more inclined with respect to the polar axis (64.8° against 55°).

If considered and exploited together, the two systems are synergic and permit increased reliability of the so-called navigation of precision.

The main technical difference concerns the carriers' frequency that in the GPS is equal for all the satellites, while in the GLONASS it is different for each constellation of satellites, incrementing of about 560 kHz.

In GLONASS and GPS, mixed systems are necessary to increase the minimum number of satellites from 4 to 5 needed to solve the 3D position of the surveyed point. This is due to the time misalignment between the two constellation clocks, which participates as a further unknown.

The visibility of the receivers is higher if both the positioning systems are used, permitting a potential availability of 48 (24+24) satellites. Consequently, as opposed to the urban domain, the positioning in hilly and mountain zones and woody areas becomes easier and more reliable. As the useful satellites are those having an elevation higher than 7.5° and simultaneously available, their number ranges from 7 to 14 in relation to the latitude and to the exposition of the receivers.

Errors to be considered in the combined use of GLONASS/GPS systems are

- errors in ephemeredes prediction;
- satellites' instability;
- clocks' instability;
- delays due to atmospheric effects: ionosphere and troposphere;
- thermal disturbances because of the receiver.

Potential accuracies using the single or combination systems with respect to the measurement strategies are shown in Table 7.5.

Table 7.5 Positioning precision of a point independently measured by the GLONASS and GPS systems and their different integrations in pseudo-distance and differential

Positioning system	Metre
GLONASS	15
GPS	10
GPS+GLONASS	3–5
GPS differential + GLONASS	0.5
GPS + GLONASS differential	1
GPS differential + GLONASS differential	0.5

A larger constellation also guarantees better performances through measurement of phase difference, reducing the time necessary to determine the phase ambiguity by more than 60%.

The combined system offers a higher *robustness* due to the fact that the transmission on different frequency bands reduces the risk of electromagnetic interference.

7.3 Galileo Global Positioning System

The purpose of Galileo is to create a worldwide navigation system to reduce, for strategic and economic reasons, the dependence of the European Union from the GPS USA system.

The size of future needs in the transportation sector and an appropriate global coverage cannot be faced by the use of a single satellite navigation system. GPS and GLONASS seem not to guarantee the reliability and the availability required especially by the transport of people. The European system Galileo would allow correction and integration of these lacks.

Galileo is a European initiative for the development of a Global Navigation Satellite System (GNSS), which can provide services of high accuracy positioning for civilian use.

The programme schedules the realization of a new generation of satellites, represented by a constellation that will provide services for satellite navigation, combined with suitable ground infrastructures, which will constitute the Galileo system. The interoperability with the GPS system, and hence the integration between the two navigation systems, will allow the development of a global system for the GNSS navigation. The existing GLONASS and other countries without their own navigation system are participating in the development of GALILEO.

As it can regularly offer binary frequencies, Galileo will guarantee a positioning accuracy of the order of 1 m in the free service, available for all users.

7.3.1 Positioning Services

The Galileo constellation will have the support of three geostationary satellites of the EGNOS (*European Geostationary Navigation Overlay Service*) network able to provide services for civilian use and to implement an alarm system in case of malfunctioning, causing a reduction of the signal integrity, GPS and GLONASS systems included. EGNOS will be able to assure the reduction of the accuracy restrictions of GLONASS. These initiatives are related to the American systems WAAS (*Wide Area Augmentation System*) and the Japanese system MSAS (*MTSAT-Multi-functional Transport Satellite–Satellite-Based Augmentation System*).

The Galileo system is planned to provide five types of positioning service (Box 7.3):

- *free service*, Open Service, OS, basic level, dedicated to general interest applications, with availability of time and positioning signals, a service which is provided free of charge. This service will have simple and cheap receivers with a few metres ground resolution. The service will be enhanced by the possibility of integrating GPS signals, with more advantages especially in the urban domain;
- *restricted access service* for professional and commercial applications, Commercial Service, CS, requiring higher performances to offer added value services: these signals are encrypted and controlled by access keys. Typical applications are radio-diffusion of data, service guarantees, precise time signals, ionospheric delays prediction model, local differential corrections of the extreme precision positioning signals;
- *service reserved* to government, police and customs applications, *Public Regulated Service* (PRS), characterized by signal high availability and continuity. Its contribution will be relevant in public security critical circumstances. The main quality of the PRS is the signal consistency with respect to any obstacle, block or fraud;

Box 7.3 Galileo positioning services

	Free (OS)	Commercial (CS)		Ruled (PRS)		Lifesecurity (SoL)	Search and rescue (SAR)
covertures	global	global	local	global	local	global	global/local
Horizontal accuracy (h)	4 m binary 15 m mono	<1 m	<10 cm signal increased locally	6.5 m	1 m signal increased locally	4–6 m frequency binary	<10 cm signal increased locally
Vertical accuracy (v)	8 m binary 35 m mono			12 m			
Availability	99.8%	99.8%		99–99.9%		99.8%	99.8%
Integrity	No	Services with added values		Yes		Yes	Yes

- *service for passengers security*, Safety of Life Service (SoL), used for transport applications especially when on board navigation systems are damaged or out of order. This service provides the same time and positioning accuracy as the OS. The main difference is the high level of integrity, without interruptions, at the global scale for applications in maritime, aerial and railway security critical situations. This service will be guaranteed using binary frequency certified receivers;
- *humanitarian Search and Rescue service* (SAR), to localize with extreme precision messages and signals wherever on the Earth, reducing the false alarm signals to a minimum. The service is realized in cooperation with three geostationary COSPAS-SARSAT satellites, in addition to four low-orbit supporting satellites (*Low Earth Orbit*).

As far as security is concerned, the system guarantees the physical protection of the basic infrastructure and provides precise signals in emergency or war conditions. Any kind of improper use of the signal or access by hostile parties in case of conflict is considered impossible.

In order to better understand the Galileo system's potentiality, it is important to introduce some definitions:

- *Availability*: the time percentage, in a definite interval of time, in which the foreseen position accuracy is lower than a specific value, for each point within the signals volume.
- *Integrity*: including three aspects:
  - *alarm limit*: the activation of error threshold before an alarm;
  - *time of alert*: the time separating the activation of the alarm from its visualization onto the user interface;
  - *integrity risk*: the probability that a horizontal or vertical calculated positioning error overcomes the corresponding alarm limit without the operator is informed within the specific time of alert.
- *Continuity*: the probability that the requirements of a specific service performance are supported by the system for the whole time interval in the surveyed area.

7.3.2 Technical Characteristics

Galileo's central system is constituted by a constellation of 30 satellites, distributed on three orbital planes, for a complete coverage of the Earth (*Mid Earth Orbit*, MEO) through circular orbits around the Earth, at 23,616 km altitude, with low technological risk and high performances. Twenty seven satellites have fixed orbits while three have free orbits. The satellites are inclined at 56° angle over the equatorial plane, enabling coverage over 75° north.

This scheme is similar to the existing positioning systems already adopted and which will be implemented by the United States and Russia for the enhancement and update respectively by the *GPS Block IIF* and *GLONASS M* series.

The Galileo project is structured to operate at five different frequencies, with 11 different types of signals, which improve the performances and the positioning accuracies provided by the GPS and GLONASS systems (Table 7.6).

E_2-L_1-E_1 and E_{5a}/L_5 are common with the GPS frequencies; E_{5b} frequency is common with GLONASS L_3 for a complete interoperability between the systems (Table 7.7).

Galileo and GPS refer to the same geodetic coordinates system, WGS84, and their on board atomic clocks are well synchronized (Table 7.8).

This determines positioning planimetric accuracy, in some configuration, up to 1 m, with the possibility to reach horizontal accuracies lower than 10 cm by increasing the signal locally. The satellites' position is monitored by a network of ground stations strategically positioned, in order to keep under control the constellation itself and the performance of the on board clocks.

Table 7.6 The Galileo system is designed to operate in five separate frequencies, with a total of 11 different types of signals

Signal			Central frequency (MHz)		Encryption	Service
No.	Name	Type		Code	Data	
1	E_{5a}-I L_5	Data	1176.45	No	None	OS/SoL
2	E_{5a}-Q L_5	Pilot	1176.45	No	–	OS/SoL
3	E_{5b}-I	Data	1207.14	No	Partial	OS/SoL/CS
4	E_{5b}-Q	Pilot	1207.14	No	–	OS/SoL/CS
5	E_6-A	Data	1278.75	Governmental	Yes	PRS
6	E_6-B	Data	1278.75	Commercial	Yes	CS
7	E_6-C	Pilot	1278.75	Commercial	–	CS
8	E_2-L_1-E_1-A	Data	1575.42	Governmental	Yes	PRS
9	E_2-L_1-E_1-B	Data	1575.42	No	Partial	OS/SoL/CS
10	E_2-L_1-E_1-C	Pilot	1575.42	No	–	OS/SoL/CS
11	L_6	Data	1544.10			SAR

Table 7.7 Frequencies of GPS, GLONASS and Galileo systems. Galileo's frequencies E_{5a}/L_5 and E_2-L_1-E_1 are the same as GPS' frequencies L_5 and L_1; Galileo's frequency E_{5b} is the same as GLONASS' frequency L_5

Central frequency (MHz)	1176.45	1207.14	1227.60	1246+k	1278.75	1544.2		1575.42		1602+k
GPS	L_5		L_2					L_1		
GLONASS		L_3		L_2						L_1
Galileo	E_{5a}/L_5	E_{5b}			E_6	L_6	E_2	L_1	E_1	
Galileo Service	OS	OS			PRS	SAR		OS		
	SoL	SoL			CS			SoL		
		CS						CS		

Table 7.8 Main characteristics of GPS, GLONASS and Galileo systems

Characteristics	GPS	GLONASS	Galileo
Frequency of the free signal	1.023 MHz	0.511 MHz	1176.45 MHz
Strategic errors induced Selective Availability (SA)	Yes	No	No
Frequency of the binary signal	10.23 MHz	–	1207.14 1176.45 MHz 1575.42 MHz
Encryption	Yes	No	PRS/CS
Carrier Frequency			
L_1 E_2-L_1-E_1	1575.42 MHz	1602 + k × 0.5625 MHz	1575.42 MHz
L_2	1227.6 MHz	1246 + k × 0.4375 MHz	
E_6	–	–	1278.75 MHz
Satellites number in the constellation	24	24	30
Orbital planes number	6	3	3
Satellites number per orbital planes	4	8	10
Orbit inclination	55°	64.8°	56°
Orbit radius	26,560 km	25,510 km	23,616 km
Orbit period	11 h 58 min	11 h 15 min	14 h 22 min
Geodetic reference	WGS84	PZ-90	WGS84

Some ground systems will provide integrations of positioning signals sent through existing communication networks, e.g. mobile telephones, or through dedicated lines, allowing accuracy improvement and data availability, continuity and integrity.

7.3.3 Applications

American GPS, Russian GLONASS and European GNSS positioning systems already allow (the first two) and will allow (the third one) many applications and uses. They are quickly and continuously increasing, both as a consequence of the GPS S/A elimination and of the new research experiences. Galileo can provide data supporting many fields of geomatics and obtain information about the positioning of users in different sectors.

For example in the field of environmental and territorial management, the data can support

- land survey for cartographic purposes;
- landslides and subsidence monitoring;
- bathymetric and geophysical survey;

- positioning of sensors for detection of air, water and land pollutions also in movement;
- survey, monitoring and testing of airport, railway and road infrastructures;
- as well as other positive applications for
- *satellite and aerial navigation*: control of the platform position;
- *road transport*: vehicle positioning, car navigation, speed control, guide systems at a distance;
- *social services*: help for disabled people;
- *judicial and custom systems*: location of controlled people, frontier monitoring;
- *search and safety*: means for people localization in the sea and in mountains;
- *management of the emergencies*: environmental, civilians and military emergencies;

and moreover, precision farming and fishing, geo-marketing, individual mobility and many others.

Each application requires the development of a methodology and the choice of the appropriate receiver. Thus according to the final application, it is convenient to define a plan that satisfies the specific needs.

The survey of the territory for cartographic purposes or object positioning often requires accuracy higher than a few decimetres. In this case, phase smoothing submetric receivers are used, which are based on *pseudo-distance* measures and use the phase measurement only as a filter. They are relatively cheap with respect to the geodetic measurement, and the acquisitions of new points generate errors that are independent from the previously acquired points, avoiding the problem of error propagation.

Sub-metric receivers are suitable for surveys and positioning in geology, in geophysical prospecting, for precision farming, but not for cadastral surveys where geodetic positioning instrumentation with precision ranging between few centimetres and millimetres is necessary, which can be obtained by very advanced technological solutions (type of antenna, etc.).

7.4 Summary

Satellite positioning systems are used for topographic–geodetic surveys and in terrestrial, aerial and sea navigation. The main global positioning systems are the NAVSTAR (USA), largely used all over the world, the GLONASS (Russia) and the planned GALILEO (Europe).

In general a satellite positioning system is constituted by three elements:

- *space segment*: the satellites' constellation; the satellites (NAVSTAR GPS and GLONASS have at least 24 satellites each, Galileo 30 in the final configuration) positioned on inclined orbits with respect to the equatorial plane. The function of this element is to transmit data to the ground station through a complex signal, keeping the time reference, receiving and executing the instructions coming from the control segment;

- *control segment*: constituted by ground stations, which track the satellites, verify the on board clocks, correct the orbits, register new data on the satellites to be transmitted to the users;
- *user segment*: constituted by the users, equipped with passive receivers which provide the coordination of points on the Earth's surface.

GPS are based on the measure of the satellite–receiver distance. This information, arriving to the receiver from at least four satellites at the same time, gives the 3D position of the receiving point. The information sent from the satellite to the ground receiver is a signal, which is transmitted in one of the following frequencies: L_1 or L_2 (NAVSTAR GPS), L_3 or L_4 (GLONASS), E_2-L_1-E_1, E_{5a}/L_5, $E_{5/b}$ (Galileo). The codes P (precision code) and C/A (coarse precision) are bit-code signals modulated on these frequencies. Once the receiver captures the signal from the satellite, it is reproduced with a delay due to the propagation in space. The satellite–receiver distance measure can be derived in two ways:

- *pseudo-distance measures*, or *code measures* (Δt): based on the time delay derived comparing the bit-sequence code (also defined pseudo-range) sent from the satellite and the one reproduced by the receiver;
- *phase measures* ($\Delta\lambda$): based on the calculation of the difference between the phase of the L_1 (and L_2) sent from the satellite and the phase of the same wave reproduced, with a delay, by the receiver.

The measure of the coordinates of a point can be absolute, relative or differential and it can be performed in a static or cinematic way. According to the frequency and the type of measure, the receivers are distinguished:

- *code* or *pseudo-distance* receivers;
- *one-frequency* receivers;
- *double-frequency* distance.

Errors in positioning can occur: errors of measurement due to the ground receiver, systematic errors, orbit errors, tropospheric and ionospheric errors or errors due to the system manager.

Points positioning of both GPS and Galileo refer to the WGS84 reference system, while GLONASS refers to the PZ90 coordinate system. Conversion of the data between the two is possible.

Galileo system is planned to provide five types of positioning service: *Open Service* OS, dedicated to general interest applications; restricted *access service* for professional and commercial applications, Commercial Service, CS; *service reserved* to government, police and customs applications, *Public Regulated Service* (PRS); *service for passengers security*, Safety of Life Service (SoL); *humanitarian Search and Rescue service* (SAR), to localize with extreme precision messages and signals wherever located on the Earth.

Further Reading

El-Rabbany A., 2006, Introduction to GPS: The Global Positioning System, 2nd ed. Artech House Publishers, Boston, p. 250, ISBN 1596930160/9781596930162.

Hofmann-Wellenhof B., Lichtenegger H., Wasle H., 2008, GNSS – Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and More. Berlin Springer, p. 516, ISBN: 978-3-211-73012-6.

Parkinson B.W., Spilker J.J. (Eds.), 1996, Global Positioning System: Theory and Practice.. American Institute of Aeronautics and Astronautics, Inc., Washington, DC, Vols. I and II.

Bibliography

Bao-Yen Tsui J., 2005, Fundamentals of Global Positioning System Receivers: A Software Approach). Series in Microwave and Optical Engineering. John Wiley & Sons Inc., New York, ISBN: 0-471-70647-7.

Departments of Defense, Homeland Security, and Transportation, 2005, Federal Radionavigation Plan. National Technical Information Service, Springfield, Virginia, USA, 22161, DOT-VNTSC-RITA-05-12/DoD-4650.5

Hoffmann-Wellenhof B., Lichtenegger B.H., Collins J., 1994, GPS: Theory and Practice, 3rd ed. Springer-Verlag, New York.

Kaplan E.D. (Ed.), 1996, Understanding GPS: Principles and Applications. Artech House Publishers, Boston.

Leick A., 2004, GPS Satellite Surveying, 3rd revised ed.. John Wiley & Sons Inc., New York, p. 435.

Steede-Terry K., 2000, Integrating GIS and the Global Positioning System. ESRI Press Redlands (CA) USA, p. 95.

van der Marel D.H., 2000, US Discontinue Intentional Degrading of GPS, GIM International, June, pp. 51–53.

Wells D. (Ed.), 1989, Guide to GPS Positioning. Canadian GPS Associates, Fredericton, NB, Canada.

Chapter 8
Digital Image Processing

Methodologies used in *Remote Sensing* (RS) directly and indirectly interact with other techniques such as image analysis and pattern recognition.

The image analysis, or *Digital Image Processing* (DIP), provides RS with the theoretical and numerical instruments for the digital processing of remotely sensed images and for their enhancement in function of the interpretation and classification. The expression *Digital Image Processing* refers to the process performed to enhance the image, in order to facilitate the extraction of information concerning objects in imagery.

The pattern recognition, instead, is mostly dependent on the interpretation skills of the expert, who also uses mathematical and statistical instruments in image interpretation, generally aimed at the classification in function of any correspondence between the objects themselves and numerical values.

Considering the huge amount of data that remote sensors acquire, the introduction of DIP automatic interpretation procedures is necessary to limit processing time maintaining high accuracy of the information derived concerning observed phenomena. Human experience and interpretation skills and the photo-interpretation techniques can be only partially replaced by automatic procedures, which can be used by the expert to better analyse and understand the contents of an image (Fig. 8.1).

Before going into detail, it is fundamental to define the term *image*, as the data acquired by remote sensors on satellite, aerial or ground platforms are available in digital format with specific spatial, radiometric and spectral characteristics.

A digital or numerical image is a two dimension regular distribution, deriving from sampling according to a grid pattern and the values that each elementary cell can assume ranging in a certain numerical interval. Hence it is a group of discrete elements organized in rows and columns defining the coordinates of the image, thus represented by a matrix. Each discrete element of the matrix, or elementary cell, called pixel (*Picture Element*), is associated with a *Digital Number* (DN). The pixel is the image element that represents the corresponding ground truth.

Digital Numbers (DNs) express a discrete measure of the *radiance* (L) detected by the sensors and measured in Watts per square metre per steradian ($W \cdot m^{-2} \cdot sr^{-1}$). Actually physical measures of the radiation measured by the *detectors* are continuously acquired and later transformed in discrete levels by analogical/digital converters. Radiance values are resampled according to a definite number of bit in function

M.A. Gomarasca, *Basics of Geomatics*, DOI 10.1007/978-1-4020-9014-1_8,

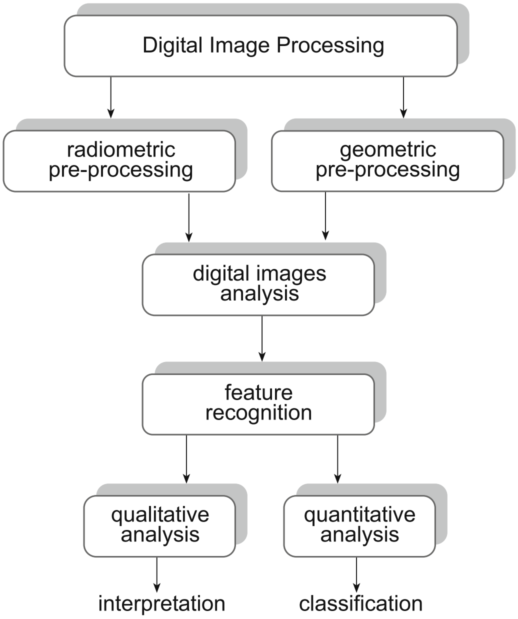


Fig. 8.1 Flow of digital image processing

of their characteristics and of the system's radiometric resolution. For example, in 8 bit resolution digital images, 256 Digital Numbers or grey levels (from 0 to 255) can be represented (Fig. 8.2).

Another important characteristic of a remotely sensed image is the *spectral resolution*, i.e. the wavelength interval (λ) to which the radiance represented by its Digital Number refers. For the same scene, several images can be available, each one

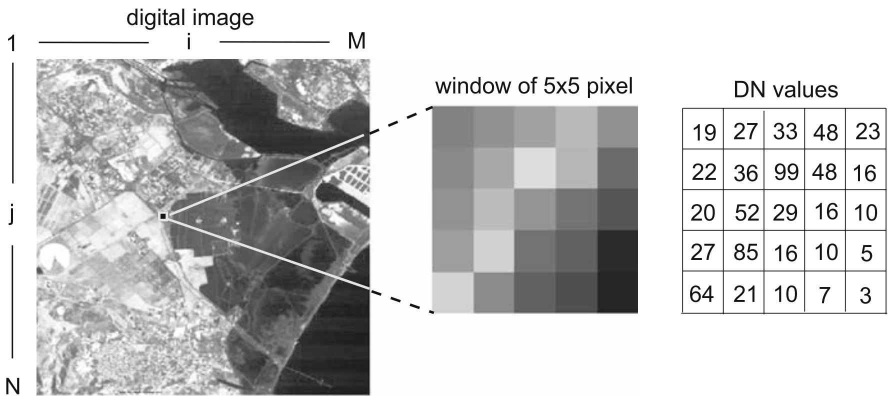


Fig. 8.2 Representation of a digital image (**a**), zoom on a 5 × 5 pixel window of the image (**b**), the correspondent Digital Numbers in the window (**c**)

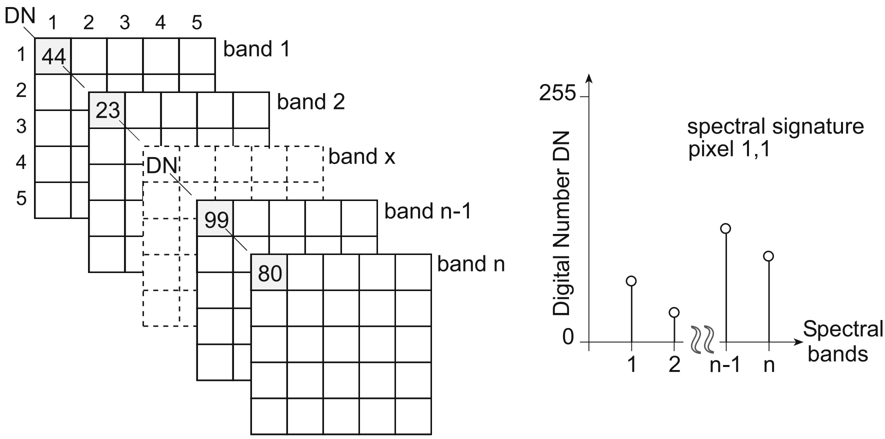


Fig. 8.3 The cells or pixels (picture element) constitute a multispectral image in a grid model. The radiometric values in Digital Numbers (DNs) of the pixel with coordinates 1,1 are reported

referring to the radiance recorded in definite spectral ranges. Each pixel is hence characterized by a value or Digital Number (DN_1, DN_2, $DN_{n\text{-}1}$, D_n) as many as the spectral bands, i.e. by an *n*-dimensional vector (Fig. 8.3) where the DN is a discrete unit whose variability range depends on the radiometric resolution.

8.1 Image Transformation

Image transformation can be on *single band* or *multiple bands* (*multispectral*). A synthetic classification is reported in Box 8.1.

Box 8.1 Classification of the image transformation

Single band	Multispectral
Radiometric correction	Algebraic operation with bands
Geometric correction	Vegetation indices (VI)
Image to image registration	Tasseled Cap
Histogram modification	Principal components analysis (PCA)
Digital filter	Classification

Single-band transformation is based on algorithms (e.g. algebraic operations) independently applied to each one of the bands acquired by a sensor. The correction procedure of radiometric errors is the main single-band transformation, where the cause is corrected for each band. In geometric corrections and image registration, each pixel of the transformed image is assigned the radiometric value corresponding to the nearest pixel in the original image (*nearest neighbour resampling*), or, more commonly, to the average of a variable number of pixels next to that one. Other

single-band transformations are the operations carried out in digital image processing techniques, such as the *modification of the histograms* enhancing the contrast in image visualization, and the *digital operators, or filters*. The operators, operating through mobile windows, assign the radiometric value of each pixel of the filtered image in relation to the original values and those of the neighbours.

In *multispectral transformations,* bands can be combined (e.g. ratios of two spectral bands) to produce, from the originals, synthetic bands, also called pseudo-bands. These transformations can generate new images with higher information (the case of Vegetation Indices and Tasseled Cap) used in the interpretation phase or as processing (the case of the Principal Components Analysis or of the algebraic operations on the bands) before an automatic classification.

8.2 Pre-processing

The data collected by sensors on board different platforms, before being used for the interpretation need to be processed to correct errors due to the noise and distortions generated during the acquisition and transmission. Some of these operations (*system pre-processing*) are carried out in the receiving stations (control and reception segment) while other residual operations (*basic pre-processing*) are left to the user's skills (user's segment). The introduction of high geometric and radiometric resolution satellite data is transferring many of the operations performed until now by the receiving stations to the skilled user. In this situation, the two operative segments (the control/reception and user) are no longer assigned rigid tasks in radiometric and geometric corrections.

8.2.1 Radiometric Pre-processing

Distortions caused by radiometric errors listed in Box 8.2 generate effects on the radiometric values of the image pixels, inducing a non-representative distribution of the spectral band brightness. This distortion, which can be eliminated or reduced by *radiometric pre-processing*, depends on:

- errors introduced by the sensors' mis-functioning during the acquisition;
- geometric characteristics of the acquisition system and Sun position, in particular its inclination measured by the zenith angle;
- atmospheric layer between the sensor and the detected scene which affects the spectral responses as well as decreases the contrast in the scenes. The atmosphere influences remotely sensed data acting as a barrier, limiting the electromagnetic wave propagation by absorption phenomena and acting as an unreal source of energy by *scattering* phenomena.

Radiometric distortions in many applications are negligible or represent the study object (e.g. meteorological applications with Meteosat data), or they are source of

Box 8.2 Radiometric and geometric effects

Error	Cause	Type of distortion
Radiometric	Sensor	Radiometric calibration of the sensor Anomalies in the scansion (line striping effect)
	Geometry of the system	Effect of the Sun angle elevation Soil inclination (leaning)
	Atmosphere	Radiation absorption (subtractive) Atmospheric diffusion (additive)
Geometric	Acquisition system	Perspective deformation (asset variation, velocity and altitude) Panoramic distortion Distortion due to the opto-mechanical scan system oscillation
	Atmosphere	Atmospheric refraction
	Earth shape	Rotation, Earth curvature and orography

noise if the extraction of physical phenomena information is the purpose (e.g. surfaces' reflectance and temperature). In these cases, the following factors should be considered:

- re-definition of the sensor's original physical unit (radiance) through calibration and correction processes of radiometric anomalies related to technical scanning problems;
- normalization of the signal based on the scene's illumination conditions and to its orography;
- signal clearing from atmospheric components.

8.2.1.1 Radiometric Effect Depending on the Sensor

Radiometric Calibration of the Sensor

The radiometric calibration transforms the image's Digital Numbers, in each spectral band, into radiance (L) values measured by the optical systems' detectors. This transformation is generally expressed by parametric linear functions related to calibration coefficients specific to each sensor and functional to the wavelength or the spectral bands. The following relation determines the physical radiance units $L_0(\lambda)$ measured in each Landsat spectral band derived from the digital signal levels (DN) acquired by the sensor.

$$L_0(\lambda) = \mathrm{DN}(\lambda) \times gain(\lambda) + offset(\lambda) \tag{8.1}$$

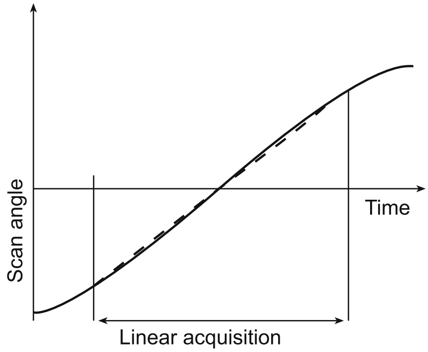


Fig. 8.4 Non-linearity of the sensor with rotating mirror scan system in function of the time. The non-linear continuity of data acquisition to the extremes limits this effect

where

$L_0(\lambda)$: spectral radiance $(W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1})$
$DN(\lambda)$: Digital Number derived from the input spectral radiance
$gain(\lambda)$: gain coefficient or system's amplification $(W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1} \cdot DN^{-1})$
$offset(\lambda)$: movement coefficient $(W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1})$
λ : wavelength or spectral range (μm)

where gain and offset coefficients are constants in relation to Landsat 5's TM, or in relation to Landsat 7 ETM+. Moreover, as the detectors are subjected to physical decay, with consequences on the remotely sensed signal, the calibration coefficients depend also on the time (Fig. 8.4). During the satellite system's operativity, it is necessary to verify the coefficients in the function (8.1) with appropriate updating of the calibration coefficients (Fig. 8.5). The definition of the calibration coefficients and their updating can be obtained by different methods:

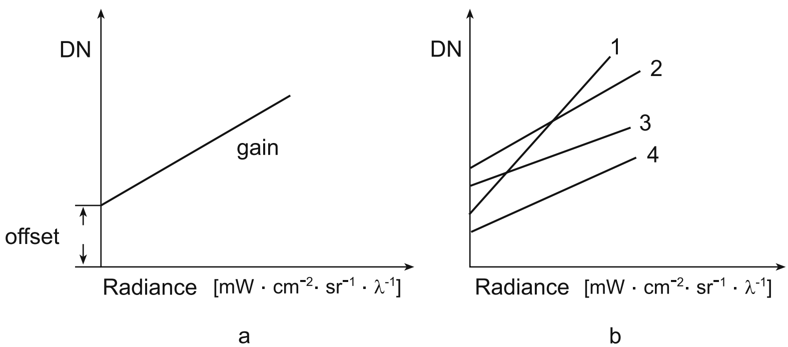


Fig. 8.5 An ideal detector has a linear relationship between the radiation and the signal (**a**). In case of radiometric mismatching among the bands 1, 2, 3 and 4 (**b**), specific corrections are requested to obtain uniform signals. The coefficient of calibration, gain and offset are used to convert the digital signal levels of a defined spectral band, in physical units of radiance

- *in field*, before the launch;
- *on board* the space platform, after the launch and *indirect* data *acquisition* during the flight;
- using ground reference elements, after the launch.

In-field calibration verifies the system before it is integrated in the space platform. The pre-launch absolute radiometric calibrations, for example of Landsat 7 ETM+, are based on measures of a characterized and calibrated spatially uniform source. The source illuminates the detector, and the back signal is compared with the standard established by the *National Institute of Standards and Technologies* (NIST), obtaining pairs of gain and offset for each spectral band of the ETM+ sensor.

Post-launch on board and *indirect* calibrations aim at monitoring in-the-time variations of the sensor's absolute radiometric reply. The main uncertainties are associated with the satellite launch, whose energy can alter the sensor's characteristics defined in the pre-launch test, but also the strong vacuum surrounding the sensor and its natural decline are factors that may affect its efficiency. The post-launch *on board calibration* of Landsat 7 ETM+ is performed by an internal calibrator made up by three lamps with different attenuation filters programmed in order to provide eight different levels of irradiance. On board radiometric calibration, having pre-launch values as reference, highlights any variations that occurred after the launch, providing a high-precision panoramic view about the sensor's behaviour over periods ranging from a few hours to months. *Indirect calibrations* provide independent tools (e.g. Moon observations) to verify the lamps' conditions over periods of months or years. Studies on 10-year data collected by the Landsat 5 TM shows that the absolute calibration of visible and near infrared spectral bands has an uncertainty lower than 5% (Thome et al., 1997).

The main aspect of *ground reference* methods is the prediction of the radiance measured by the sensor above the atmosphere over a sample area that is compared with remotely sensed data. In order to assess the radiance in situ measures, synchronous to the passing of the satellite, and radiative transfer codes, needed to study the behaviour of the radiation through the atmosphere, are used. In the approach based on reflectance measures, the ground data refers to reflectance measures of standard surfaces and to the characteristics of the atmosphere above the sample area. In the approach based on the radiance, a calibrated radiometer on board an airplane measures the radiance of a sample area and the atmosphere's optical properties. As in the reflectance approach, these data are used to determine, by the equations of radiative transfer, the radiance expected to reach the satellite's sensors (Slater et al., 1987, 1996). Both in reflectance and radiance measures, the calibration site, including *ground reference,* has to consider the sensor's geometric resolution characteristics as well as the variability due to meteorological conditions.

Scan Anomalies (Striping, Break Lines, Skip Pixel)

Sensors constituted by a group of detectors can generate regular *striping* effect.

In optical–mechanical scanning acquisition systems during a single oscillation of the mirror (*whiskbroom* system) or for each scan line, the signal is measured by

a group of many detectors (e.g. 6 for the Landsat MSS and 16 for the Landsat TM and ETM+). Due to temperature variations and possible modifications of supports and materials, single detectors' transfer functions are not identical with one another, causing striping effect of lines transversal to the direction of the satellite's orbit.

Also the *pushbroom* acquisition systems, where single strips are simultaneously acquired by a high number of detectors, can generate inter-calibration defects. In these images, the striping effect is along the columns, i.e. along the direction of the satellite's orbit.

In both cases, striping effects are more evident on homogeneous areas (snow, sea and lake water, desert) (Fig. 8.6). The noise can be reduced by calibrated ancillary data or by statistical techniques. In statistical techniques, the image data are used to obtain a relative correction of the sub-images of each single detector, based on the hypothesis that over a homogeneous and wide enough area each detector is exposed to the radiance of the scene with the same probability distribution. The method consists in the equalization of the mean μ_i and of the standard deviation σ_i (Box 8.5) of the data acquired by the detector *i* over the image according to the following linear function:

$$DN_c = DN_{nc} \cdot \frac{\sigma}{\sigma_i} + \mu - \mu_i \cdot \frac{\mu}{\mu_i} \tag{8.2}$$

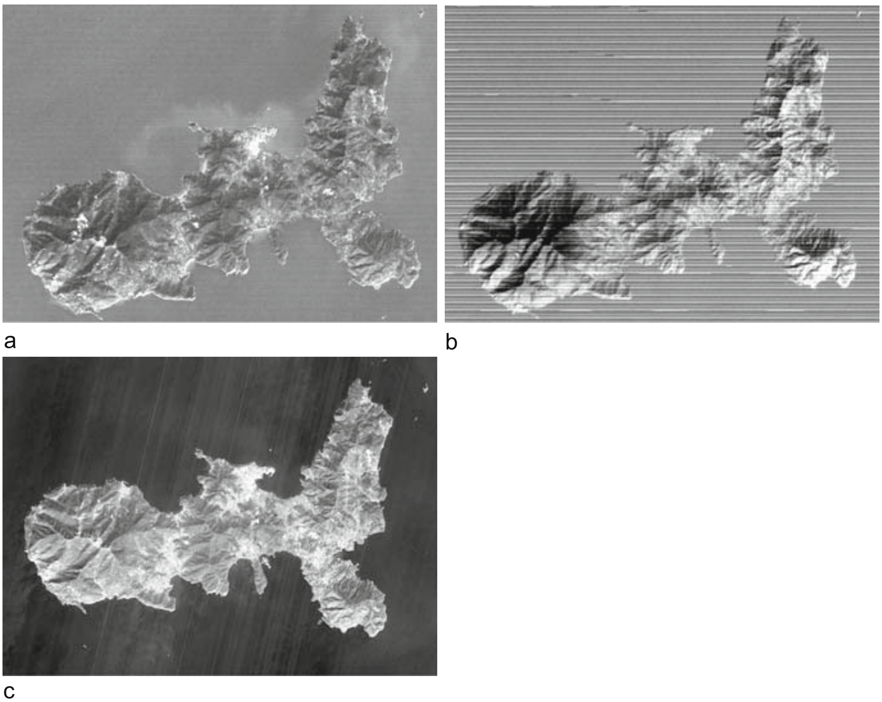


Fig. 8.6 Sea-striping effect surrounding Elba Island, Tyrrhenian Sea, Italy, of Landsat TM1, blue band 0.45–0.52 μm; a: correct sensor calibration; b: striping due to the loss of signal by the system; c: radiometric error of the sensor LISS-III of the satellite IRS-1C, 5.6 m ground resolution and 6 bit of dynamic interval; the strips are not vertical because the image is geometrically corrected

where

DN_c : corrected Digital Number
DN_{nc}: not corrected Digital Number
μ_i σ_i : arithmetic mean and standard deviation relative to the detector's lines
μ and σ : respectively arithmetic mean and standard deviation of the whole image.

Out-of-scale values as break line and drop line and skip pixels are other possible scan anomalies, usually due to the system's tension falls or signal loss.

In these cases, simple correction techniques are used:

- lines with out-of-scale values are replaced by previous or following line, or by their average;
- isolated pixels with out-of-scale values are replaced with the average of the nearest pixels (3×3 or 5×5 windows).

The obtained Digital Numbers are artificial and after insertion among original data the image's pixels values can be normalized.

8.2.1.2 Radiometric Effects Depending on the System Geometry

The geometric configuration of the Sun-surface-sensor system during the acquisition affects the response of the surfaces in quantitative terms, acting on the radiant fluxes and modifying both the intensity and the shape.

Sun Angle Effect

In multi-temporal acquisitions for *change detection* analysis or mosaicing of many scenes' image, data should be normalized according to the Sun angle as reported in (8.3). This correction, based on a flat surface pattern, does not eliminate the shadows effect due to the orography or the atmospheric effects.

Terrain Inclination (Slope)

The incident radiant flux density per surface unit, i.e. the irradiance, does not depend only on the Sun angle but also on the relative orientation of the surface with respect to the direction from which the light comes. The irradiance is highest when the direct irradiance propagation direction is orthogonal to the surface and, with regard to the light, is lowest when the ground is in the shadow. In this condition, the direct incident irradiance over the surface is null, and only the irradiance component scattered by the atmosphere is present.

When the image's pixels do not reproduce horizontal surface, i.e. the Sun zenith angle does not coincide with the angle between the Sun and the surface, the relative position between the Sun and the surface must be calculated by the Digital Terrain Model (DTM) of the study area, in order to retrieve the pixels lying in terms of the three director cosines.

The effects of the Sun angle and of the ground inclination, i.e. of the geometry between source (Sun), surface and sensor, can be expressed by the *Bi-directional Reflectance Distribution Function* (BRDF). This function describes the anisotropic behaviour of the radiation reflected by surfaces that only ideally have a *Lambertian* behaviour, i.e. isotropic. The measure of the radiance reflected by the surfaces in fact depends on the relative position between sensor and source (i.e.: colour variation of objects observed when the Sun is in front or behind). Data acquired by remote sensing sensors, in particular those with a wide swath, are affected by natural surfaces' non-Lambertian behaviour. These effects can represent noise as well as a source of information. In the first case, some algorithms normalizing the radiance values with high zenith observation angles into nadir observations will be used; in the second case, multidirectional information (same surface seen from different view angles) allows acquisition of more data for the estimation of the surfaces' physical parameters (e.g. albedo).

8.2.1.3 Atmospheric Effects

The term atmosphere means the gas shell surrounding the Earth and, due to its complex chemical, physical and dynamic actions, affecting most of the phenomena taking place on our planet's surface. Not least is the interaction with the radiation: atmospheric gases, aerosols and clouds scatter, absorb and refract the electromagnetic waves modulating the signal that reaches the sensor and alter the original radiance reflected/emitted by the observed surface. All these phenomena, called the *atmospheric effect*, can modify the surface observations by remote sensors, requiring ad hoc correction in order to obtain realistic interpretation of the final image.

Atmospheric Absorption and Scattering

Atmospheric gases, aerosol and vapours contribute to scatter, absorb and refract the incident solar irradiance and the radiance reflected and/or emitted by the terrestrial surface. The scattering spreads the energy changing its propagation direction; the spectral absorption reduces the quantity of energy available in a certain wavelength, while the refraction modifies only the radiation path direction. The global effect is a reduction of the contrast among the observed objects.

The atmospheric scattering is a physical phenomenon occurring when the radiation interacts with the particles composing the atmosphere (from the molecules to the aerosols). Scattering phenomena, explained by Rayleigh and Mie Laws (see Chapter 4), cause a multi-directional re-irradiation. Radiance values measured by the sensors are produced by the sum of surface radiance and by atmosphere radiance reflected towards the sensor, which makes a considerable contribution, especially in the shorter wavelengths.

The other physical phenomenon of interaction between the radiation and the atmospheric particles is the radiation absorption at given wavelengths. The absorption is due to the presence of water vapour, ozone and carbon dioxide which, acting as selective filters, makes the atmosphere opaque in some portions of the

electromagnetic spectrum called atmospheric windows, characterized by high transmission of the radiation. The atmospheric absorptions thus restrict the observations (passive remote sensing) to the atmospheric windows where the radiation absorption is lower.

8.2.2 Atmospheric Correction

In different applications (e.g. determination of ground values of reflectance, albedo and temperature; extraction of biophysical parameters; some applications of change detection), it is necessary to correct the images from the atmospheric effects. This correction, besides clearing the signal from atmospheric absorption and scattering effects, also includes the radiometric calibration and the normalization of the effects due to the system geometry. In practice, through the atmospheric correction, the main radiometric effects in remotely sensed images are normalized.

There is no unique method for the atmospheric effect correction that is simple, accurate and widely used. As a consequence, numerous methods have been developed, for specific kinds of problems and different levels of accuracy. Based on the quantity and on the accuracy of atmospheric parameters needed to apply them, the methods for studying atmospheric effects can be distinguished into two groups:

- models based on the physics of the radiative transfer;
- methods based on the images (image based).

The first category includes the models solving the equation of electromagnetic energy radiative transfer through the atmosphere (*Radiative Transfer Code*). To accurately describe the radiation propagation, these models need in situ collection measures of the optical properties of the atmosphere acquired at the same time as the scene acquisition. In general radiative transfer codes combined with in situ atmospheric measures produce the most accurate assessments. Nevertheless, all these procedures are often too expensive and complicated to be commonly applied. The biggest disadvantage of this kind of correction is that it requires measures of atmospheric parameters simultaneously with each passing of the sensor. This is often impossible in monitoring programmes, or when historical series have to be analysed.

In order to overcome this inconvenience, other correction methods have been proposed, where the information about the atmospheric properties can be retrieved from the image. As it is easy to understand, these models (*image based*) produce less accurate results compared to radiative transfer models, however they are widely used and are a good alternative when the atmospheric properties during the image acquisition are not known. Among the most known image-based methods, the *dark-pixel* (or *dark-subtraction* or *haze-reduction*) method is based on the hypothesis that in the scene at least one pixel has a reflectance equal to zero. This pixel's radiance contribution hence depends only on the atmospheric component which is subtracted from all the image's pixels. Among the surfaces that better reply to the dark-pixel hypothesis, there are deep and oligotrophic water bodies or sharp shadowed areas.

Summarizing, the radiance measured by the sensor is the result of complex energy balances involving the illuminated surfaces and the atmospheric layer between them and the sensor. The atmosphere influences the balance by additive (atmospheric scattering) and subtractive (absorption) components according to the diagram of Fig. 4.18 (see Chapter 4.3.6: Electromagnetic Radiation–Atmosphere Interaction).

In the applications of the *Surfaces Remote Sensing*, whose purpose is to assess the surface's spectral reflectance $\rho_{\mathrm{g}}(\lambda)$ from the radiance $\mathrm{L}_0(\lambda)$ measured by the sensor above the atmosphere, these components have to be recognized in some way, quantified and subtracted to the signal, in order to reconstruct the effective radiance of the surfaces.

In the applications of the *Atmospheric Remote Sensing*, aiming at studying the atmospheric properties (e.g. applications for meteorology), the atmospheric effects, and in particular absorptions, are the signal to be investigated and analysed.

Concerning surface Remote Sensing, the equations on which the atmospheric correction procedures are based will be dealt with shortly.

8.2.2.1 Computation of the Surface Reflectance

According to the hypothesis of *Lambertian* surface, where the incident radiation is homogeneously or isotropically reflected by the surface in all the directions, the recorded signal can be corrected from the effects due to the illumination conditions according to the following relation:

$$\rho_a(\lambda) = \frac{\pi \, L_0(\lambda)}{E_0(\lambda) \cos \theta_z} \tag{8.3}$$

where

$\rho_a(\lambda)$: apparent reflectance (as perceived by the sensor)
$\mathrm{L_o}(\lambda)$: radiance measured by the sensor above the atmosphere ($\mathrm{W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}}$)
$\mathrm{E_o}(\lambda)$: solar irradiance above the atmosphere incident on a surface perpendicular to the radiation ($\mathrm{W \cdot m^{-2} \cdot \mu m^{-1}}$). $\mathrm{E_o}(\lambda)$ includes d^2 considering the effect of the distance *d* Earth–Sun during the scenes' acquisition (d changes in the year within the range 0.983–1.017)
θ_z : incidence angle of the solar flux directed towards the Earth surface, defined as Sun zenith angle.

To obtain ground reflectance values $\rho_{\mathrm{g}}(\lambda)$ also the atmospheric effects have to be considered, quantified through the values of radiance backscattered by the atmosphere due to scattering phenomena $\mathrm{L_{atm}}(\lambda)$ and by coefficients of atmospheric transmittance $\tau(\lambda)$ (values ranging from 0 to 1), including the loss of energy due to *absorption* and *scattering phenomena*. The atmospheric transmittance, besides being a function of λ and of the atmospheric particles' concentrations, also depends on the thickness of the atmosphere (optical path): when the Sun's rays arrive with a certain inclination θ, $\tau(\lambda)$ decreases, as the light signal has to cover a longer distance. According to Moran et al. (1992), the surface reflectance can be calculated

by the following relation that is a simplification of the radiative transfer equations (for example, multiple reflection of near pixels' contribution is omitted)

$$\rho_g(\lambda) = \frac{\pi[L_0(\lambda) - L_{d\uparrow}(\lambda)]}{\tau_{v\uparrow}(\lambda)[E_0(\lambda)\cos\theta_z\tau_{z\downarrow}(\lambda) + E_{d\downarrow}(\lambda)]} \tag{8.4}$$

where

$L_o(\lambda)$: radiance measured by the sensor above the atmosphere ($W\cdot m^{-2}\cdot sr^{-1}\cdot\mu m^{-1}$)
$L_{atm}(\lambda)$: atmospheric radiance backscattered in direction of the sensor's field of view ($W\cdot m^{-2}\cdot sr^{-1}\cdot\mu m^{-1}$)
$E_o(\lambda)$: solar irradiance above the atmosphere ($W\cdot m^{-2}\cdot\mu m^{-1}$). $E_o(\lambda)$ includes d^2 considering the effect of the distance Earth–Sun, while the product [$E_o(\lambda)$ $cos\theta_z\tau$ (λ, z)] represents the solar irradiance directed to the surface
$E_d(\lambda)$: scattered solar irradiance descending to the surface due to the scattering of the solar radiation flux ($W\cdot m^{-2}\cdot\mu m^{-1}$) in the atmosphere
τ (λ, ν) : atmospheric transmittance along the optical path surface-sensor
τ (λ, z) : atmospheric transmittance along the optical path Sun-surface.

Resolving Eq. (8.4) for the superficial spectral reflectance $\rho_g(\lambda)$, the unknown parameters are

- atmospheric spectral radiance scattered up $L_{atm}(\lambda)$;
- atmospheric irradiance scattered to the surface $E_d(\lambda)$;
- atmospheric transmittance scattered up and down, respectively $\tau(\lambda, \nu)$ and $\tau(\lambda, z)$,

while the known terms are

- solar radiance values $L_0(\lambda)$ can be obtained from the image's Digital Numbers for each band using the sensor's radiometric calibration coefficients;
- solar irradiance values $E_0(\lambda)$ are constant except for d^2 (8.3) and depend only on the reply functions of the sensor's filters (for Landsat 5, for instance, they were published by Markham and Barker in 1986) or appropriate width squared filters can be admitted;
- zenith angle θ can be calculated from the date, the position and the time of the satellite passage.

The unknown quantities in Eq. (8.4) can be determined by radiative transfer models which, from inputs (atmospheric visibility, climatic model, aerosol model, etc.) calculate, besides the others, the atmosphere radiance, the ascending and descending transmittance and the irradiance scattered by the surface. More precisely the radiative transfer codes implement more complex formulations of Eq. (1.4), which can also include the effects of BRDF and of the ground inclination, aiming at determining the more likely value of the superficial spectral reflectance $\rho_g(\lambda)$. Instead, using the dark-pixel method, the ascending and descending transmittances are supposed equal to 1 (or $\tau(\lambda, \nu)=1$ and $\tau(\lambda, z)=\cos\theta_z$), the scattered irradiance null (($E_d(\lambda)=0$) and the radiance scattered by the atmosphere $L_{atm}(\lambda)$ equal to the satellite radiance $L_o(\lambda)$ read in correspondence of the dark-pixel.

8.2.2.2 Computation of the Surface Temperature

The density of the radiant flux emitted by the surfaces (exitance) is a function, according to Plank Law, of the surfaces temperatures. Similarly, to determine the blackbody temperature of the ground surfaces from the radiance measures, it is enough to invert the Plank Eq. (4.6) and, knowing the emissivity of the surfaces, it is possible to determine their real temperatures. The spectral domain characteristic of the emitted energy's Remote Sensing applications is the so-called thermal infrared where, according to the Earth's mean temperatures, electromagnetic emission phenomena became relevant.

The detection of the radiation emitted by the terrestrial surface is also influenced by the atmosphere layer that it crosses. The temperature of a surface detected by a thermal sensor hence can be higher or lower than the one measured on the ground according to the emission or the absorption by the atmospheric components prevailing over the energy emitted by the detected surface. Similarly to what was seen in the determination of the surfaces' reflectance, the radiance recorded by a sensor operating in the thermal infrared domain and set on a satellite can be enounced as follows:

$$L_0(\lambda) = \tau(\lambda,\mathrm{v}) \cdot \varepsilon(\lambda) \cdot \mathrm{L_{sup}}(\lambda) + \tau(\lambda,v) \cdot [1 - \varepsilon(\lambda)] \cdot \mathrm{L_d}(\lambda) + \mathrm{L_{atm}}(\lambda) \qquad (8.5)$$

where

$\mathrm{L_0}(\lambda)$: radiance measured by the sensor ($\mathrm{W{\cdot}m^{-2}{\cdot}sr^{-1}{\cdot}\ \mu m^{-1}}$)
τ (λ, v) : atmospheric transmittance along the surface-sensor optical path
ϵ (λ) : emissivity of the detected surface
$\mathrm{L_{sup}}(\lambda)$: blackbody radiance emitted by the surface ($\mathrm{W{\cdot}m^{-2}{\cdot}sr^{-1}{\cdot}\ \mu m^{-1}}$)
$\mathrm{L_d}(\lambda)$: descending radiance of the sky integrated over the whole hemisphere above the surface ($\mathrm{W{\cdot}m^{-2}{\cdot}sr^{-1}{\cdot}\ \mu m^{-1}}$)
$\mathrm{L_{atm}}(\lambda)$: atmospheric ascending radiance along the air column from the surface to the sensor ($\mathrm{W{\cdot}m^{-2}{\cdot}sr^{-1}{\cdot}\ \mu m^{-1}}$)

When resolving the (8.5):

- $\mathrm{L_0}(\lambda)$ can be obtained by the image's Digital Numbers using the sensors' radiometric calibration coefficients;
- emissivity ϵ, parameter known for the main surfaces;
- descending radiance $\mathrm{L_d}(\lambda)$ for a clear sky in the band 10.44–12.42 μm can be calculated knowing the absolute air temperature on the ground.

The interaction between the infrared radiation and the atmosphere is more complex, described by the variables' ascending atmospheric radiance $\mathrm{L_{atm}}(\lambda)$ and transmittance τ (λ, v), due to the difficulty of knowing the structure and optical properties of the atmosphere during the satellite acquisition. As in the determination of the surfaces' reflectance, these variables can be defined by the radiative transfer models based on the input atmospheric/climatic parameters.

8.2.3 Geometric Pre-processing

These corrections are made on the images in order to re-establish, according to the user's needs and the sensor's characteristics, the geometric correspondence between the image cells (image points) and the physical points of the observed ground area (object points). In the following treatment, the traditional approach is conserved, according to which parametric corrections of the distortions related to the acquisition geometry are supposed to be made in control and receiving station, or segment. The users, receiving pre-elaborated images from the distributor, carry out only corrections of residual distortions related to the registration and non-parametric ortho-projection operations.

The situation is more complex in the case of aircraft image acquisition as the expert user has to provide geometric correction himself.

8.2.3.1 Distortions Depending on the Platform Altitude, Velocity and Attitude

Altitude variations of a satellite platform produce scale variations, maintaining same Instantaneous Field of View (IFOV) and Field of View (FOV) constant. If the satellite's speed changes, the scale in the direction of the acquisition is modified.

The images obtained by sensors on board aircraft are subject to stronger distortion than satellites in relation to the platform's attitude instability, due to rolling, pitch, rotation, yaw and considerations about the relative altitude in percent (%) with respect to the acquisition altitude.

Aircraft are susceptible to atmospheric phenomena, such as descendent currents, causing abrupt altitude and general attitude variations, producing distortions in the image. Geometrical distortions of airborne image are shown in Fig. 8.7.

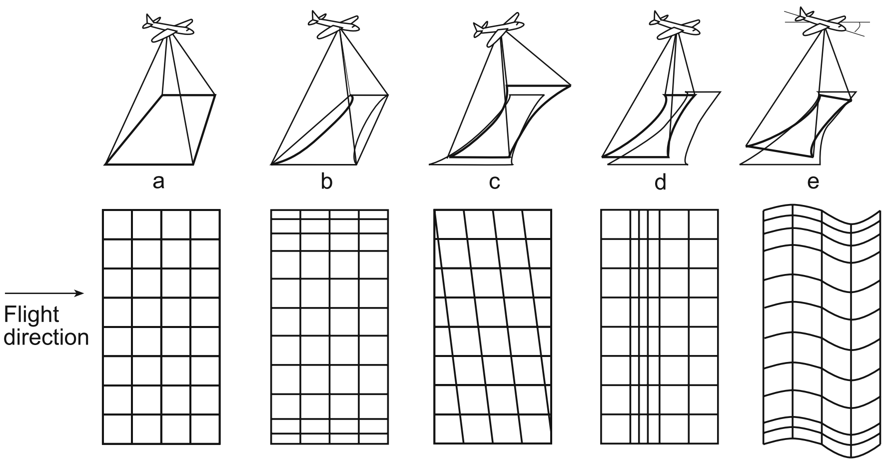


Fig. 8.7 Geometrical distortion of the image due to the platform motion (airborne or satellite). (**a**) Image with no distortion; (**b**) across-track scan to the flight line, (**c**) roll, (**d**) pitch, (**e**) yaw

8.2.3.2 Panoramic Distortion

Scanners used for spatial acquisitions have constant instantaneous Field of View (IFOV); in optical–mechanical rotating mirror sensors, the sampling time is constant, while in the array sensors the sensors' sizes are constant.

In any other acquisition system, the IFOV includes larger ground areas as going far from the vertical (nadir), hence there is information loss by compression and distortion of the relation between the ground pixel and the image pixel reproducing a *panoramic effect* (Fig. 8.8).

This kind of error is more accentuated in the case of off-nadir acquisition: very wide total Field of View (FOV) sensors are the most sensitive to this error. NOAA-AVHRR satellite (FOV: 54°) and LARA-MIVIS airborne system (FOV 70°) are strongly influenced; Landsat 7 ETM+, with FOV 7.5° (see Chapter 6), is less influenced.

These distortions are more evident in linear elements, especially if crossing the acquisition scene, such as roads, railway lines, with S-shaped distortions at the borders, due to compression effects along the scanning line. Non-linear elements undergo less evident distortions.

8.2.3.3 Distortions Depending on the Earth's Rotation

Rotating mirror optical–mechanical scanning systems with East–West swath acquisition, such as Landsat MSS, TM and ETM+ and NOAA-AVHRR, and the array

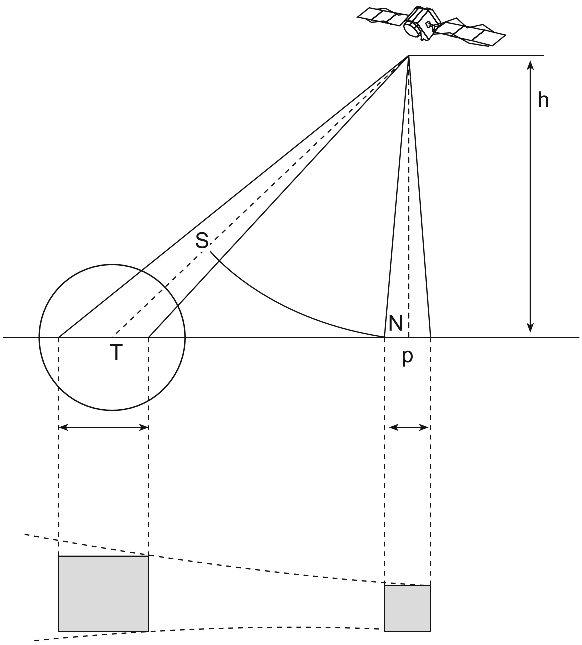


Fig. 8.8 Effect of the scan angle on the pixel size, with constant IFOV

electronic scanning systems with simultaneous acquisition along the swath width of SPOT-HRV, IRS-LISS, Terra-MODIS and ASTER, Envisat-MERIS, etc. require a definite time to acquire data in each scanning line.

The satellite track acquisition is along the swath in a north–south direction determining a distortion as the Earth rotates from west to east, and therefore it is necessary to progressively scale down the scanning lines in relation to the ground movement. The scanning lines' shifting towards the west depends on the satellite's and Earth's relative speeds and on the acquired image length:

$$\Delta x = \nu_T \cdot t \cdot \cos\alpha \tag{8.6}$$

where

ν_T : Earth's speed in m/s
t : scanning time of one scene in seconds
cos α : mean latitude of the scene to be acquired

For example, TM needs 25 s to record a 185 km × 185 km scene and at 45° latitude (cosα: 0.707) Earth speed is 463.82 m/s; the difference from the beginning to the end of the acquisition is 8200 m equal to 2°33'. The correction is made scaling down pixels towards the west as reported in Fig. 8.9.

Optical–mechanical scanning systems also have non-linear scanning towards the extremes due to the mirrors slowing down when they have to change acquisition direction (Fig. 8.5).

8.2.3.4 Distortions Depending on the Earth's Curvature

These distortions characterize satellite systems acquiring wide scenes, due to the low geometric resolution, and orbiting at high altitude like Envisat-MERIS,

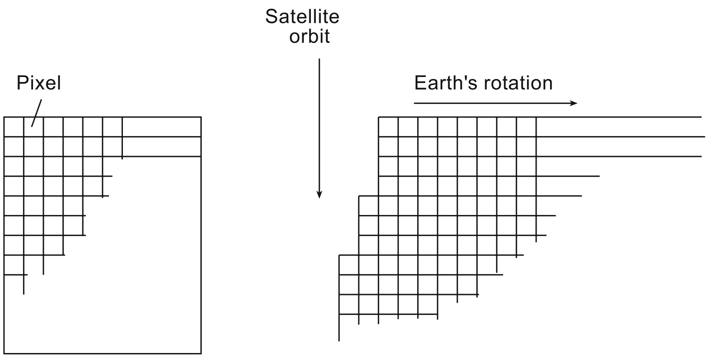


Fig. 8.9 The rotation effect causes a dislocation of sequential Western lines to compensate for the terrestrial movement during the scene's acquisition

SPOT-Vegetation, NOAA-AVHRR. At lower altitudes or at higher geometric resolution, as for many commercial satellites, the terrestrial curvature does not affect the distortions.

At the borders of the swath, the acquisition area is larger than in the vertical, or nadir, view. In NOAA imagery, covering 2,700 km^2 acquired from 833 km altitude, a pixel at the border of the scene has an area 2.89 times larger if the terrestrial curvature effect is ignored and only the panoramic distortion effect is considered, and area 4.94 times the pixel at nadir if considering the curvature (Fig. 8.10).

8.2.4 Correction of the Geometric Distortion

Transformations that modify the image geometry but not its aspect are defined as geometric corrections and give as a result the registration of the image itself. The purpose is to obtain the overlap of reference image to another image or map covering the same area.

The production of a new recorded or geometrically correct digital image requires a resampling operation generating an alteration of the image's original radiometry.

Finally the images need orthoprojection corrections to eliminate, or better decrease, the errors due to orographic reliefs (Plates 8.1, 8.2, 8.3 and 8.4).

8.2.4.1 Image Registration

Multi-temporal comparison of remotely sensed data acquired by different sensors and from different positions, defining a common reference system, is essential for pattern recognition.

The purpose is in realizing a proper geometric correspondence:

- *image to image registration:* two images of the same area acquired in different times;
- *map to image georeferencing:* image and reference topographic map.

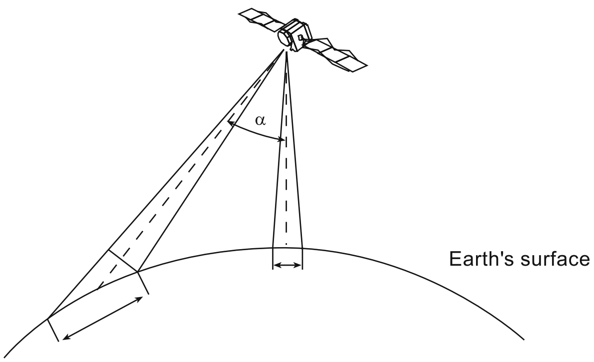


Fig. 8.10 Effect of the Earth's curvature on the pixel size to the scan direction

In the case of *image to image registration,* one image is chosen as a reference and the correspondence among homologous pixels of the two scenes generates the internal image reference.

In the case of *map to image georeferencing,* the image is adapted, after complex geometric–mathematic operations, to the cartographic reference system (Fig. 8.11). The processing that creates a correspondence between remotely sensed images and cartography is defined as *georeferencing*. Depending on the positioning accuracy and on the area's morphology, the georeferencing can be carried out by more or less complex plane transformations, i.e. *straightening*, or by orthoprojection (parametric and non-parametric) procedures, i.e. 3D transformations (see Chapter 3).

Any geometric transformation is performed in two phases:

- identification of the *mathematical–geometrical model of the transformation* and assessment of the parameters of this model (analysis). The transformation model is chosen according to the positioning accuracy to be obtained. Area orography influences the choice. Transformation models may be simple plane roto-translations with isotropic scale variation, polynomial functions of different degree, plane or 3D polynomial ratios functions. The model's parameters assessment is performed by the collimation of Ground Control Points (GCPs) whose number is a function of the number of model's parameters to be assessed. The GCPs are particular points recognized as homologous on the two considered elements (image to image, image to map) which establish the geometric correspondence between the two reference systems involved. GCPs must be chosen in correspondence of stable-in-the-time objects such as crossroads, bridges, permanent features, lines;
- *application (synthesis) of the structured model* using a radiometric resampling procedure for assigning radiometric tones to the new transformed image.

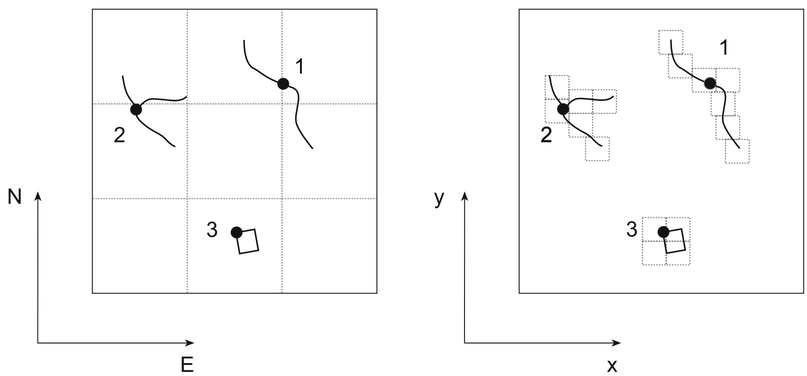


Fig. 8.11 Image georeferencing based on a defined coordinate system using ground control points (GCPs). N, E: geographic coordinates; x,y: image coordinates

8.2.4.2 Resampling

The geometric transformation process needs, as a consequence, to re-assign the radiometric value of each pixel of the new image. In the resampling phase, the original radiometry values are altered.

Resampling techniques are performed by mathematic interpolation procedures.

There are three commonly used resampling algorithms (Fig. 8.12):

- *nearest neighbour*, the new Digital Number is given by the Digital Number of the pixel having line and row coordinates nearest to the X, Y (real) coordinates obtained by the geometric transformation;
- *bi-linear*, the new value is defined by the weighted mean of the values with a surrounding of the four nearest pixels, hence modifying the original radiance value;
- *cubic*, the new Digital Number is assigned basing on evaluation of the 16 nearest pixels.

The nearest neighbour has the advantage of assigning in the new image pixel values really observed in the original image. This method is used in classification processing after resampling.

Nevertheless the digital image processing basic rule of performing a classification before any image transformation, when possible, is valid.

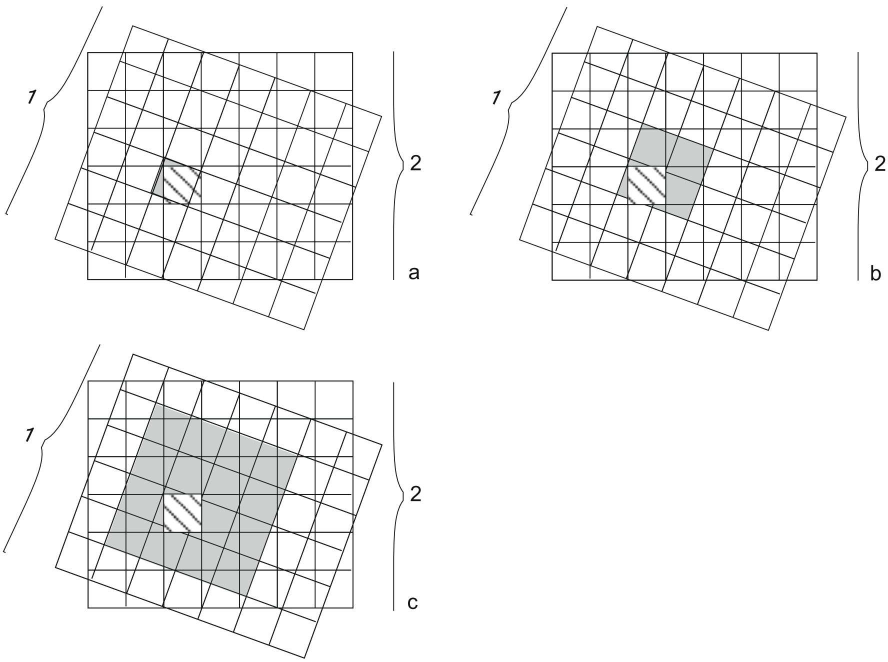


Fig. 8.12 Resampling with the three algorithms: (**a**) nearest neighbour, (**b**) bilinear and (**c**) bicubic

8.2.4.3 Orographic Effect and Orthoprojection

The presence of orographic complexity induces displacement of the image cells with respect to their correct position (the one that they should have in case of orthogonal projection on the cartographic plane).

Simple rectification procedures of modification of the plane coordinates often do not solve this problem. More complex operations, based on the integration of planimetric and altimetric data (Digital Elevation Model) and in some cases of satellites' orbital parameters, produce a more precise correction defined *orthoprojection* or orthorectification.

In the ideal case in which the sensor's position and orientation (external orientation) and internal geometry (internal orientation) are known during the acquisition, it is possible to position the image cells on the theoretical surface (Fig. 8.13). In the real case, these data are not known, and the problem is solved through 3D nonparametric transformations generally based on polynomial ratios functions (*Rational Function Model*, RFM, see Chapter 3).

In Fig. 8.14, the P point is visualized where the P' point should be: an inconsistency between the image and the truth comes up, to be corrected by the application of proper algorithms.

Geometrical distortions are also related to the platform's attitude variations and to variations in the mirror's rotation speed in the optical–mechanical oscillating sensors.

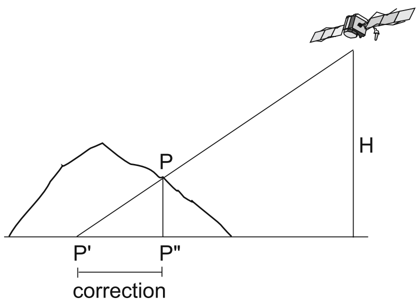


Fig. 8.13 Transformation of the projection from central (P') to orthogonal (P") for the correct positioning of the point detected by the remote sensor

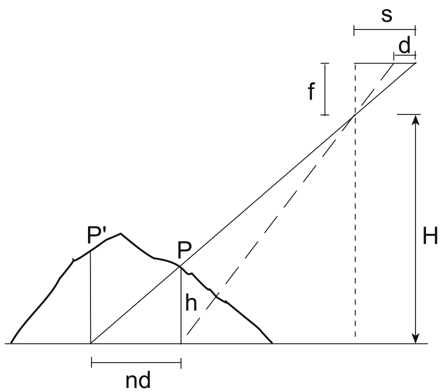


Fig. 8.14 Displacement error (d) of a point (P) beside the original scenes due to the ground morphology (orography)

Graphical error is important among the positioning errors; the thickness of the writing instrument of maps (0.2 mm is the value generally considered) can alter point positioning. This error value represents a limit threshold for the precision that can be reached at different scales: on a 1:100,000 map, it is equal to 20 m, on 1:5000 map to 1 m ground truth.

8.3 Digital Image Processing

Remotely sensed digital images can be studied by *qualitative* and *quantitative* analyses for the extraction of information.

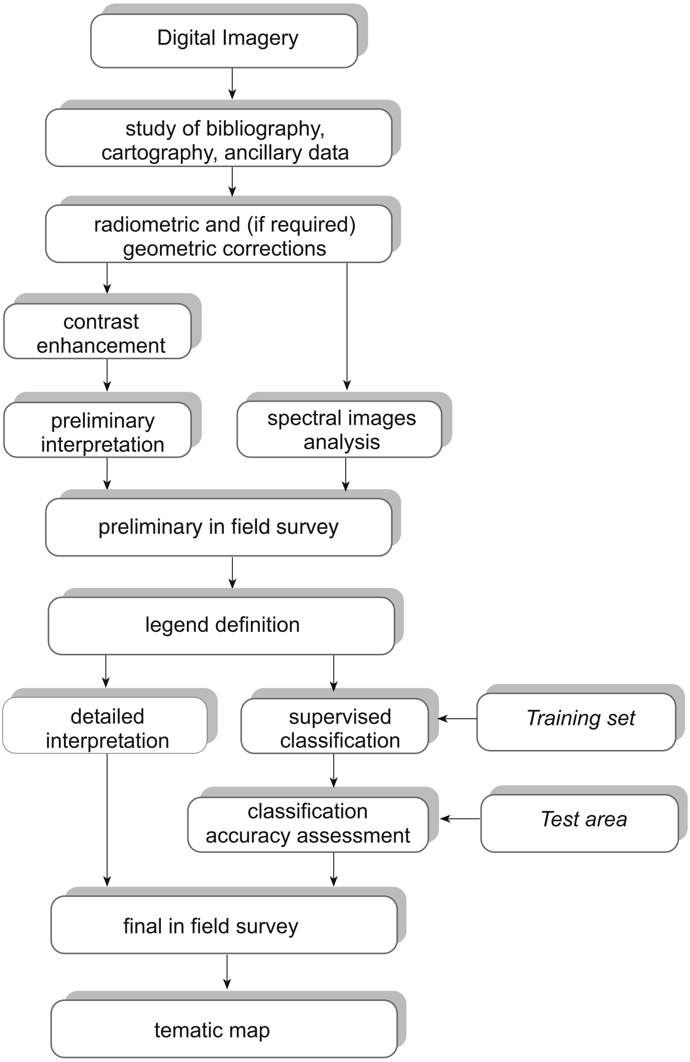


Fig. 8.15 Qualitative and quantitative (semi-automatic) digital image processing flow

These two processing procedures have considerable differences, related, firstly, to the human mind's logical mechanisms and, secondly, to computing systems' important contribution.

The qualitative analysis is related to human interpretation skills (visual interpretation or photo-interpretation), the quantitative one to the automatic (digital processing) or semi-automatic interpretation (if there are controls by the operator in one or more phases of the processing). In general the term quantitative analysis is associated with the classification process (Fig. 8.15).

The main differences between the two systems are shown in Box 8.3.

Box 8.3 Qualitative and quantitative analysis approache

Interpretation (photo-interpretation)	Digital processing
Operator: person	Operator: computer
Subjective, synthetic	Objective, analytic
Concrete	Abstract
Decision process partially dependent from the spectral behaviour	Decision process dependent from the spectral behaviour
Multispectral analysis limited at 3 bands colour composite	Multispectral analysis applicable in the multi-dimensional space
Human limitation in the grey tones and colours separation	Extension to the use of all the available grey levels (dynamic range)
Easy use of the spatial information (synoptic vision)	Spatial information hardly suitable
Operative, in general, on small areas	Operative on large areas
Operative on set of pixel	Operative on the single pixel or per fields
Low accuracy in the area surface computation	Computation of the area based on the pixel count
Long time in the interpretation process in large areas	Reduced computing timing

Before qualitative and quantitative analyses, spectral images can be processed improving visualization and facilitating the reading and interpretation by the operator (Box 8.4).

8.3.1 Spectral Analysis Techniques

The study of spectral images can be performed operating on a single pixel with punctual transformations or involving more pixels, i.e. carrying out spatial transformations.

Box 8.4 Techniques of spectral analysis of the images

Study of the spectral images		
Transformation tables	Look-Up Tables (LUT)	
Histogram		
Scatterogram		
Radiance profile		
Techniques of visualization	Density slicing	
Techniques of enhancement	Operation on the single pixel (graphs and tables)	Spatial transformation (imagery and tables)

Punctual transformations, or radiometric operations, of the single pixel modify the radiometry or Digital Number associated to that pixel. Single pixel transformation can be represented as graphs and tables.

Spatial transformations produce modification and enhancement of geometric details operating directly on the Digital Numbers of the image. Geometric enhancement techniques are characterized by operations involving the nearest pixels as well as the pixel to be transformed.

8.3.1.1 Look-Up Tables (LUT)

One of the available operations is made by transformation tables or Look-Up Tables (LUT); they help in studying spectral images facilitating the visualization. They are inserted in the hardware between the refresh memory and the analogical-digital converter, whose output signal modulates the screen's electronic brush intensity. Through the LUT image, RGB colours can be modified. The transformation modifies the correspondent values in the look-up table.

The LUT is a vector functioning as a filter: the Digital Number of the input pixel defines the position of the vector element, and the output is the DN contained in that position of the vector. For example, the LUT Red with 256 elements with highest value (255) and LUT Green and Blue with null value (0) any image appears red (Fig. 8.16).

The use of these tables allows contrast operations to be performed in very short times (milliseconds) preserving the original data, ensuring an effective interactivity between man and processor.

8.3.1.2 Image Histogram

The histogram is the simplest instrument, but at the same time the most important and useful, in digital image processing.

The image histogram is a distribution function providing, for each DN the number of pixels having that value: the image DNs are represented on the x-coordinate and the absolute frequency is reported on the y-coordinate.

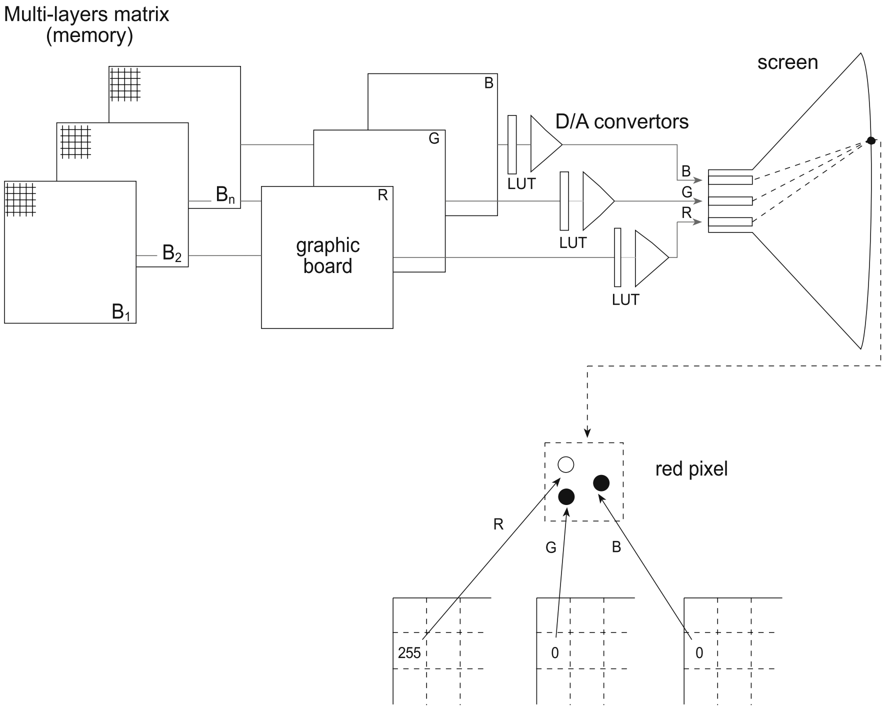


Fig. 8.16 The colour visualization principle of a multispectral image, using the primary colour tern RGB: Red, Green and Blue, on the basis of the additive synthesis. LUT: Look-Up Table, transfer functions of the input/output Digital Number

Indicating with N the number of pixels of the image and with n_x the number of image pixels with x value (where x: 0, 1, 2 ... 254, 255) the histogram function d_x is obtained associating to each value × its frequency expressed in percent:

$$d_x = \frac{n_x}{\mathrm{N}}(\%) \tag{8.7}$$

The cumulative frequencies are often represented in the histogram, too (Fig. 8.17).

The histogram characterizes the image just in statistical terms without providing information about the spatial distribution of the image's grey levels. The concept can be reversed stating that a certain histogram can correspond to more different images. Contrast enhancement techniques, described later, are based on the study of the histogram. The histogram is bi-modal or n-modal if there are two or n relative maxima, corresponding to two or n groups of objects with similar behaviour (Fig. 8.18).

Together with the histogram other statistical parameters summarize the image's grey levels content (Box 8.5):

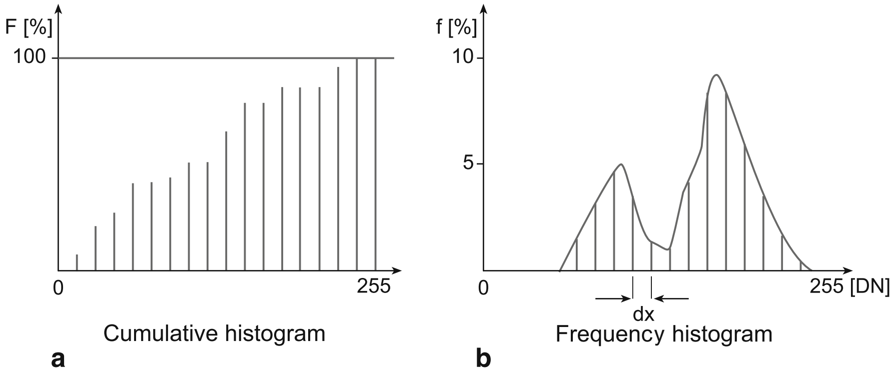


Fig. 8.17 (**a**) Histogram with cumulative frequency (F); (**b**) frequency histogram (f) of the pixels for each of the intervals d_x representing the scenes in the chosen spectral band

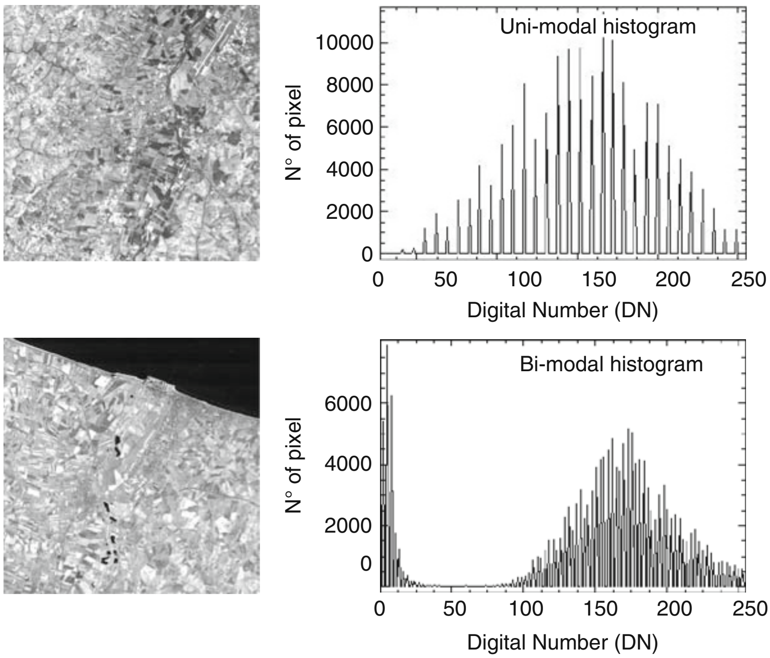


Fig. 8.18 Mono- and bi- modal histograms of the Landsat band TM4, near infrared, 0.76–0.90 μm, of two images; in the bi-modal histogram the first peak is related to the water, the second to vegetation, soil, bare soil, urban areas, etc.

- *arithmetic mean*, refers to the histogram's central tendency, corresponding to the mean brightness of the image;
- *mode*, i.e. the event occurring with higher frequency;
- *median*, the value occupying the central position in a group of data;
- *variance* and closely related *standard deviation* are measures of how dispersed a distribution is. The variance is computed as the average squared deviation of

each number from its mean. The variance analysis observes the variations in the radiance levels to understand if the image contrast is high, intermediate or low;
- *standard deviation*, square root of the variance. It is the most commonly used measure of dispersion of values around their mean value (arithmetic mean).

Box 8.5 Arithmetic mean, variance and standard deviation are statistical parameters used for a preliminary study of digital images

Standard deviation (σ) of a probability distribution, random variable, or population or multiset of values	In probability and statistics, it is a measure of the spread of its values
Standard deviation (σ)	Defined as the square root of the variance, or root mean square (RMS) deviation of values from their arithmetic mean SD measures the spread of data about the mean, measured in the same units as the data
Variance	Average of the squared differences between data points and the mean Variance is tabulated in units squared
Arithmetic mean (μ_k) of the image in band k	$\mu_k = \frac{1}{N}\sum_{ij} DN_{ij}^k$
Variance or average squared deviations of the mean (σ^2)	$\sigma_k^2 = \frac{1}{N-1}\sum_{ij}(DN_{ij}^k - \mu_k)^2$
Standard deviation or average deviations of the mean (σ_k)	$\sigma_k = \sqrt{\frac{1}{N-1}\sum_{ij}(DN_{ij}^k - \mu_k)^2}$

8.3.1.3 Scatterogram

As Remote Sensing deals with multispectral images, the scatterogram, defined as a bi- or multi-dimensional histogram, is used beside the histogram. With reference to the bi-dimensional case, to facilitate the exposition, the x-, y-Cartesian axes represent the Digital Numbers of the two considered bands 1 and 2, while the z-axis (the depth dimension) represents the frequency of occurrence of a certain phenomenon and so the number of pixels that in band 1 have the same Digital Number DN_1, defined by the correspondent x-coordinate, and in band 2 Digital Number DN_2, defined by the correspondent y-coordinate (Figs. 8.19, 8.20 and 8.21). A 3D example is reported in Fig. 8.22.

The scatterogram analyses the spectral separability of the objects in the observed scene: when clusters which take shape in a bi- or multi-dimensional scatterogram intersect there is a partial spectral separability of the objects (Fig. 8.20). The statistical parameters characterizing the scatterogram are

- arithmetical means vector of clusters;
- covariance matrix;
- coefficient of correlation among the spectral bands.

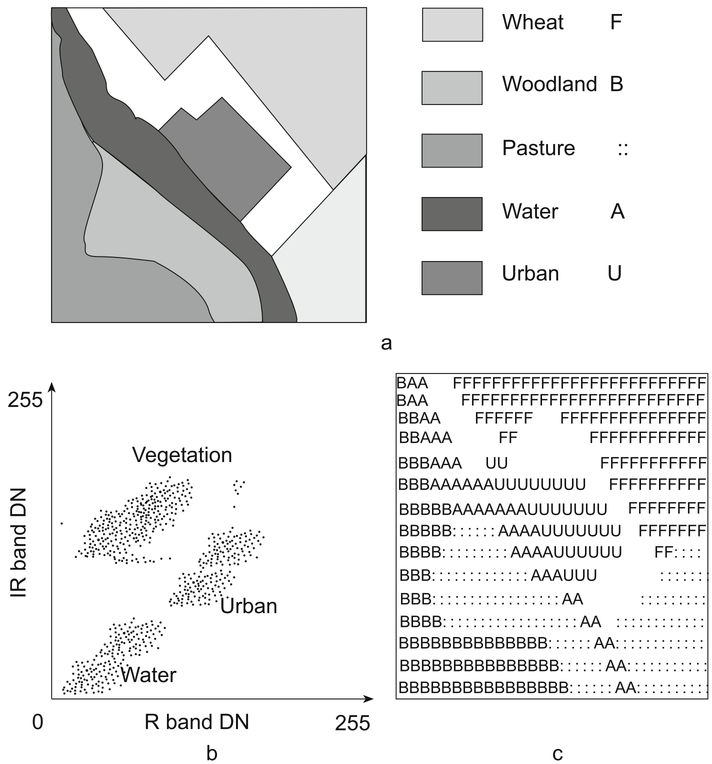


Fig. 8.19 Study of an image with its classification (**a**); bi-dimensional scatterogram of Infrared (IR), y-axis, and Red (R), x-axis, bands (**b**); image's (a) alpha-numerical rendering (**c**). DN: Digital Number

Covariance matrix is symmetric matrix whose values along the diagonal are squares and thus always positive as representing the variance of an image with respect to itself.

Covariance matrix's non-diagonal values, ranging from −1 to +1, are related to the analysis of the coefficient of correlation. It is directly dependent when the value extends to 1, indirectly when it extends to −1 and independent when the value is null.

8.3.1.4 Radiance Profile

Radiance profiles are used to explore the distribution of the pixels' values along defined paths: horizontal, vertical, oblique, along the borders.

The radiance profile describes the sequential pixel in the transet drawn in the image; the coordinates M, N of the pixels encountered along the chosen path are indicated on the x-axis, while the DNs assumed by the pixels are indicated on the y-axis (Fig. 8.23). Different representation scales are possible, absolute with DN values (0–255) or percentage (0–100).

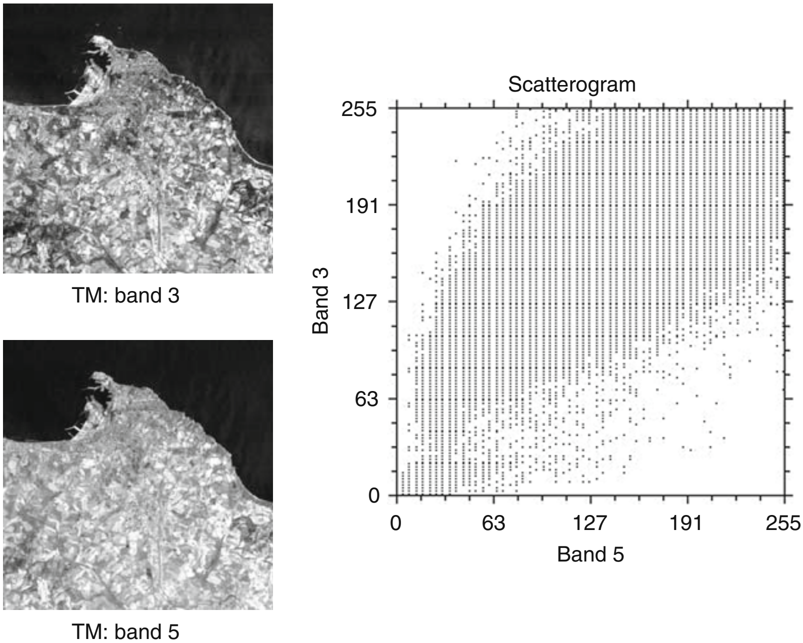
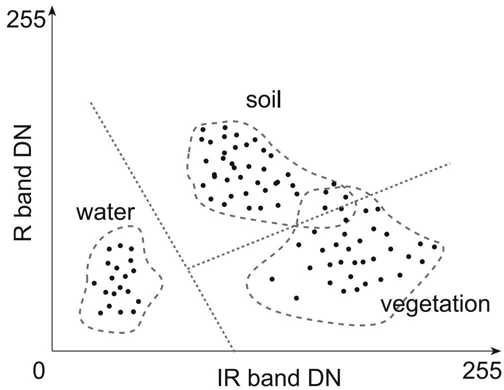


Fig. 8.20 Bi-dimensional scatterogram of two bands Landsat Thematic Mapper: band TM3, Red, 0.63–0.69 μm and TM5 1.55–1.75 μm, medium infrared

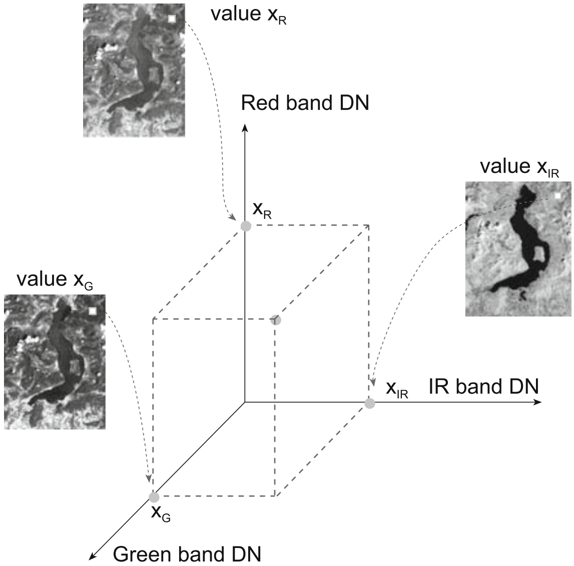
Fig. 8.21 Bi-dimensional scatterogram, with the same inverted bands on the axis with respect to Fig. 8.19b; the highest DN values in the IR correspond to the vegetation, the lowest are associated with water. In the red band both vegetation and water have low values



8.3.1.5 Density Slicing

There are different possibilities in visualizsing images for enhancing its aspect and facilitating the interpretation. Among them, the density slicing technique of visualization splits the tones' continuous scale of a spectral band, original (e.g. R band) or synthetic (e.g. Vegetation Index), in groups, each with an assigned new Digital Number. Each of the intervals so defined can be assigned with a colour, using colour palettes.

Fig. 8.22 Three-dimensional scatterogram



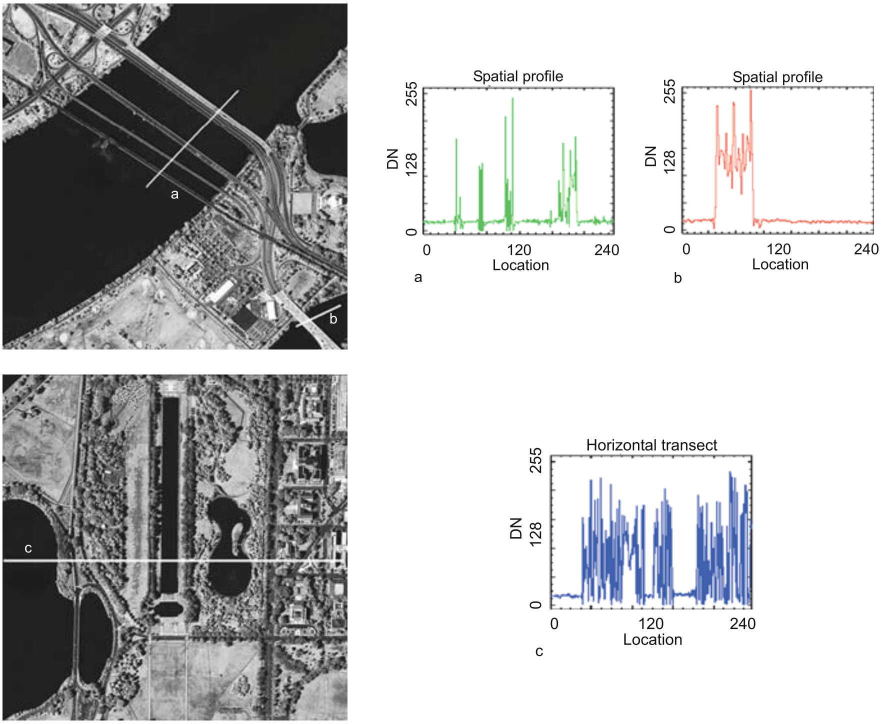
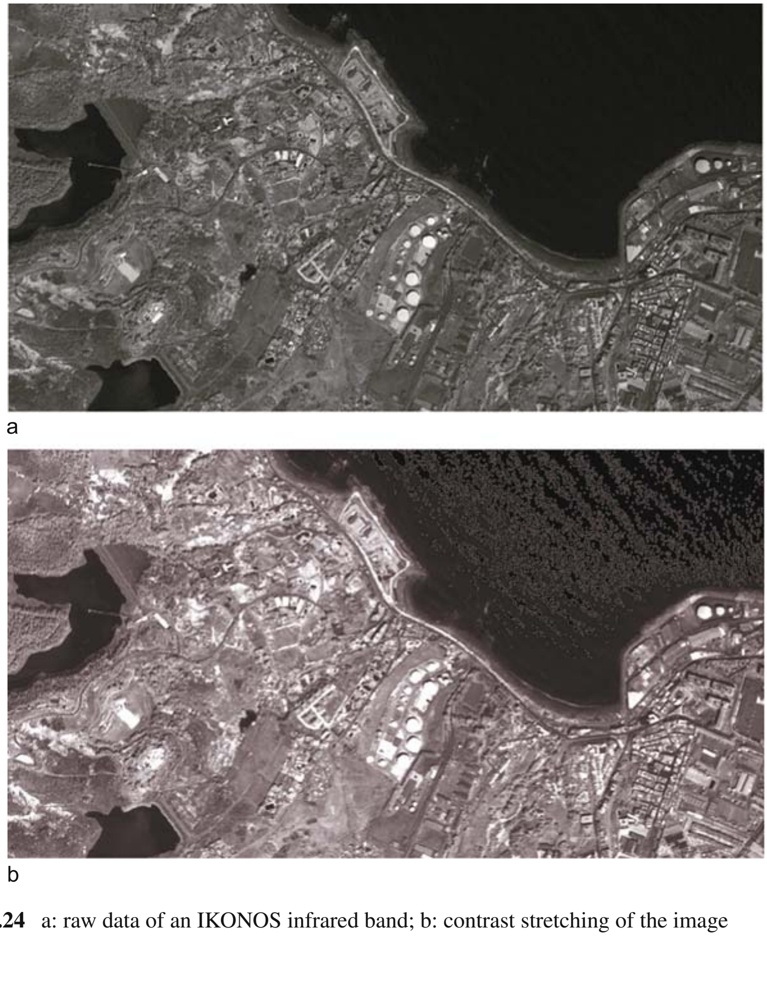


Fig. 8.23 Radiance profiles (**a, c**) of IKONOS panchromatic images 0.45–0.90 µm; b: radiance profile of a 40 pixel transect of a near infrared image, showing (W–E) in sequence a harbour inland, the dock, the sea

The assigned colours do not have direct correspondence with the true original colours, as the colour code is arbitrary as the interval of Digital Numbers. The image is visualized with codified colours or *pseudo-colour*. The *density slicing* technique associates spatial information contained in the image with the potential classes of objects in the scene helping the operator in the interpretation. For example, it is useful for a rough separation among water, soil and vegetation and, in the thermal infrared, providing the distribution of the isotherms.

8.3.1.6 Contrast Enhancement Techniques

Contrast enhancement techniques assist the interpretation in terms of parametric characteristics of the image such as contrast, texture, shape and colour (Fig. 8.24). Parameters to be used are chosen according to the specific problem to solve.



a

b

Fig. 8.24 a: raw data of an IKONOS infrared band; b: contrast stretching of the image

These techniques focus on the problem of the images' interpretation in order to facilitate the extraction of information not evident in the original images; therefore non-relevant characteristics can be eliminated or attenuated and elements of interest are highlighted (Box 8.6).

Box 8.6 Classification of the techniques for contrast enhancement of the images

Contrast enhancement techniques		
Radiometric operations on the single pixel		Geometric or spatial transformation
Linear contrast stretching	Not-uniform stretching	Filter (operator)
Simple With saturation Piecewise Cyclic enhancement	Logarithmic Exponential Equalization Gaussian	

The study of the histogram of the image provides initial information about the most important parameter of the image: the *contrast*.

The human eye can distinguish from 20 to 30 grey levels from black to white, if they are adjacent and well separated and if the operator is skilled; the contrast enhancement facilitates the perception of grey variations of an image.

The best images to be processed should have little saturation effects and well-distributed values in the available scale 0–255. Nevertheless, images might have distribution ranging in a narrow interval of values, generating weak contrast. This happens because remote sensing systems record radiance values of objects (water, vegetation, soil, rocks, glaciers, etc.) with very different spectral behaviour one from the other. Sensors are calibrated to provide signals limiting the possibilities of the saturation phenomena.

Histograms cannot be modified during the acquisition, but only during the representation: this contrast enhancement procedure is also defined as contrast accentuation or *stretching*. The algorithm is a transformation function which extends (stretches) the interval of grey levels recorded in the original image up to cover the whole dynamic range of the visualization instrument (e.g. 64: 6 bit, 256: 8 bit, 1024: 10 bit).

There are different types of contrast enhancement, and the most suitable can be determined only in function of the original image's Digital Numbers and according to the specific needs. These techniques are strongly dependent on the operator's skills.

The re-distribution of grey levels can be linear (uniform re-distribution) or non-linear (non-uniform distribution).

Linear Contrast Stretching

- *Simple*: this procedure uniformly extends the DN interval from the available minimum and the maximum values (Fig. 8.25).

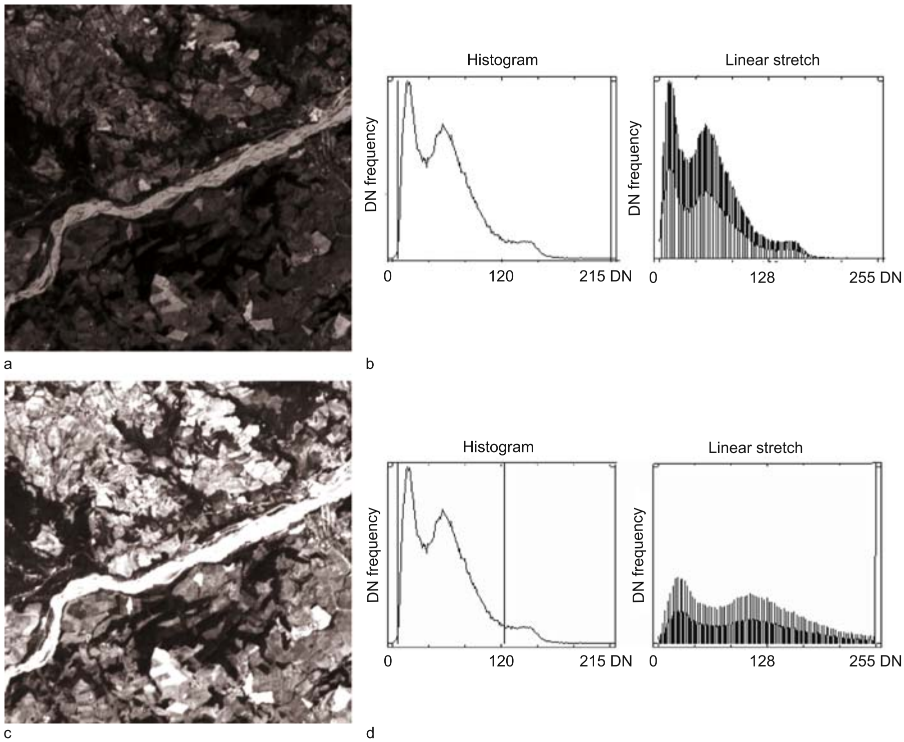


Fig. 8.25 Linear contrast stretching; a: raw data, b: histogram of the raw image a, c: linear contrast stretching, d: histogram of image in c with the interval (DN 37–126) stretched on the whole reference scale of values in 8 bit (DN 0–255)

Each pixel's Digital Number is transformed according to the relation:

$$DN_{(o)} = \left[DN_{(i)} - DN_{MIN}\right] \frac{255}{DN_{MAX} - DN_{MIN}} \tag{8.8}$$

where

$DN_{(o)}$: Digital Number of the transformed image
$DN_{(i)}$: Digital Number of the input image
DN_{MAX}, DN_{MIN} : minimum and maximum Digital Numbers of the original image (histogram's extremes)
255 : maximum value corresponding to the white in 8 bit scale (0–255).

The linear contrast enhancement does not modify the histogram shape and preserves its radiance relations.

- *Saturation*: DN_{MAX}, DN_{MIN} values can be chosen by the user obtaining the histogram saturation effect, assigning the value 0 to the $DN_{(i)} < DN_{MIN}$ and 255

to the $DN_{(i)} > DN_{MAX}$; saturation results of the order of 0.5–1% of the image's pixels at each of the extremes are considered acceptable. A simple application is for the conversion of 6 bit (0–63) or 7 bit (0–127) data into 8 bit scale (0–255). The effect of such a transformation is to increase the contrast with respect to the simple linear enhancement. Some attention must be paid to preserve the image's important structures weighting the thresholds.

- *Linear piecewise enhancement* is a particular case of contrast linear enhancement. The transformation function is no longer a straight line defined by only one angular coefficient, but a broken line, i.e. segments each with their own slope. Linear piecewise increases the control of the contrast modification process, by a succession of different expansions or compressions, offering a flexible instrument to skilled users to improve the remotely sensed images' qualitative analysis.
- *Contrast cyclical enhancement*: also called sinusoidal, is performed by dividing the available DN interval into sub-intervals, each one of which is expanded by linear re-distribution from 0 to 255 and from 255 to 0. It is used to highlight little differences within homogeneous areas, such as wooded areas, water bodies, bare soil surfaces, etc. As the new DNs are not univocally assigned, but more values are assigned to the same new value, the transformation is not reversible. Extreme use of cycles can introduce different types of noise in the image, reducing the quality.

Non-uniform Contrast Stretching

Non-linear or non-uniform transformations belong to another group of transformation functions and do not have the property of preserving the radiance ratios (Figs. 8.26 and 8.27):

- *logarithmic function*: suitable to highlight in alternative light or dark tones of an image. Usually used to enhance the contrast in dark areas, corresponding to lower digital values;
- *exponential function*: used when the radiometric values are mostly distributed around the light tones (histogram's highest values);
- *equalization*: the most used histogram transformation procedure. The image values are transformed into the new DN based on their occurrence frequency, so that a higher number of levels are used to represent the portion of the histogram with higher frequency. The percentage and cumulative frequency must be considered and sampled with intervals defined by the number of available levels in the visualization device;
- *Gaussian*: Gaussian contrast enhancement tends to real data distribution and the image histogram assumes normal shape. The re-distribution can be realized imposing that the original median of the histogram is transformed into the median of the output values, usually 127. The Gaussian transformation effect is the contrast enhancement of the image's darker areas.

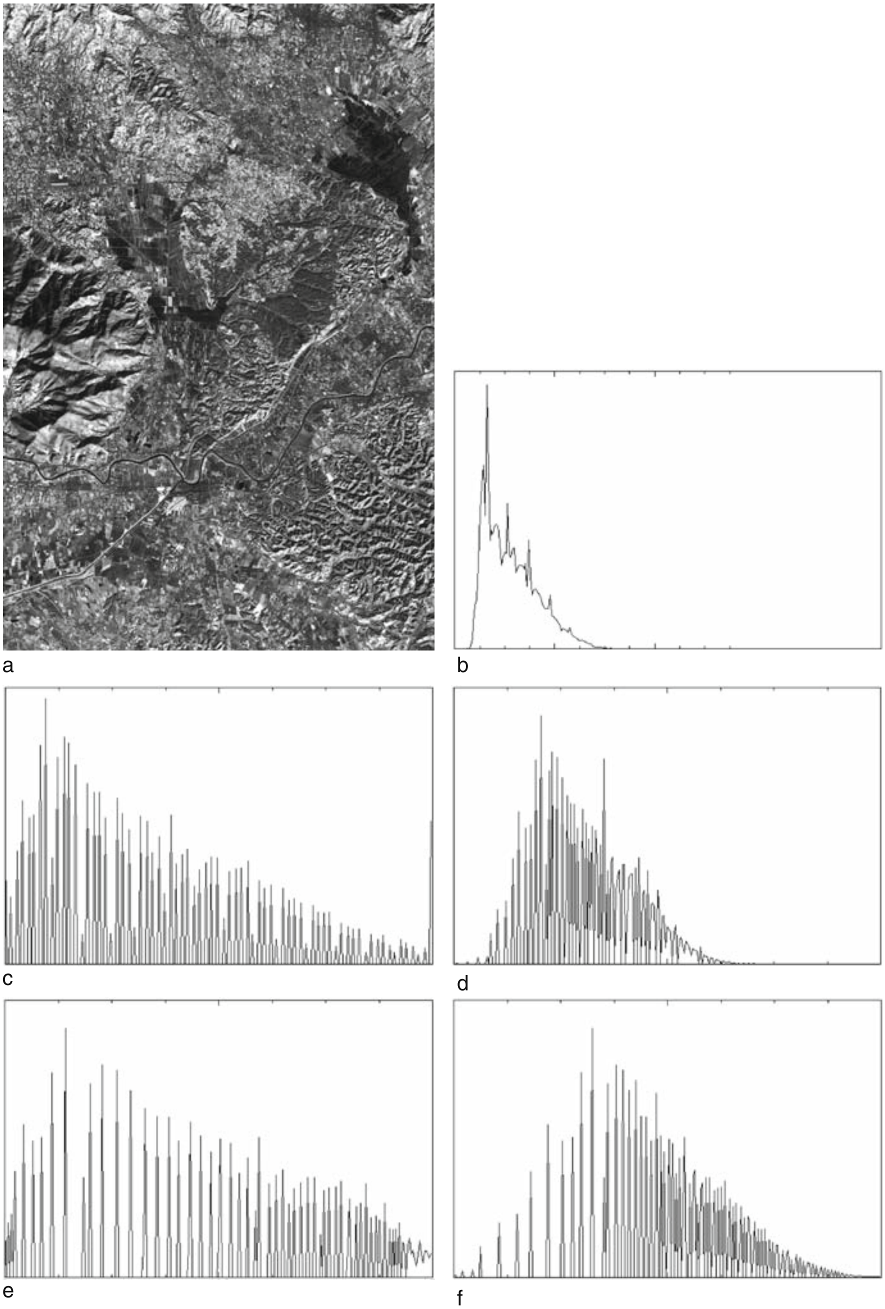


Fig. 8.26 Different types of contrast stretching on an image having a histogram with Gaussian distribution of the data. (**a**) Y-axis: percent frequency (%), left, and the number of pixels per each DN, right; x-axis: radiometric values with different contrast stretch; (**b**) original histogram; (**c**) linear contrast stretching; (**d**) exponential function; (**e**) equalization; (**f**) Gaussian

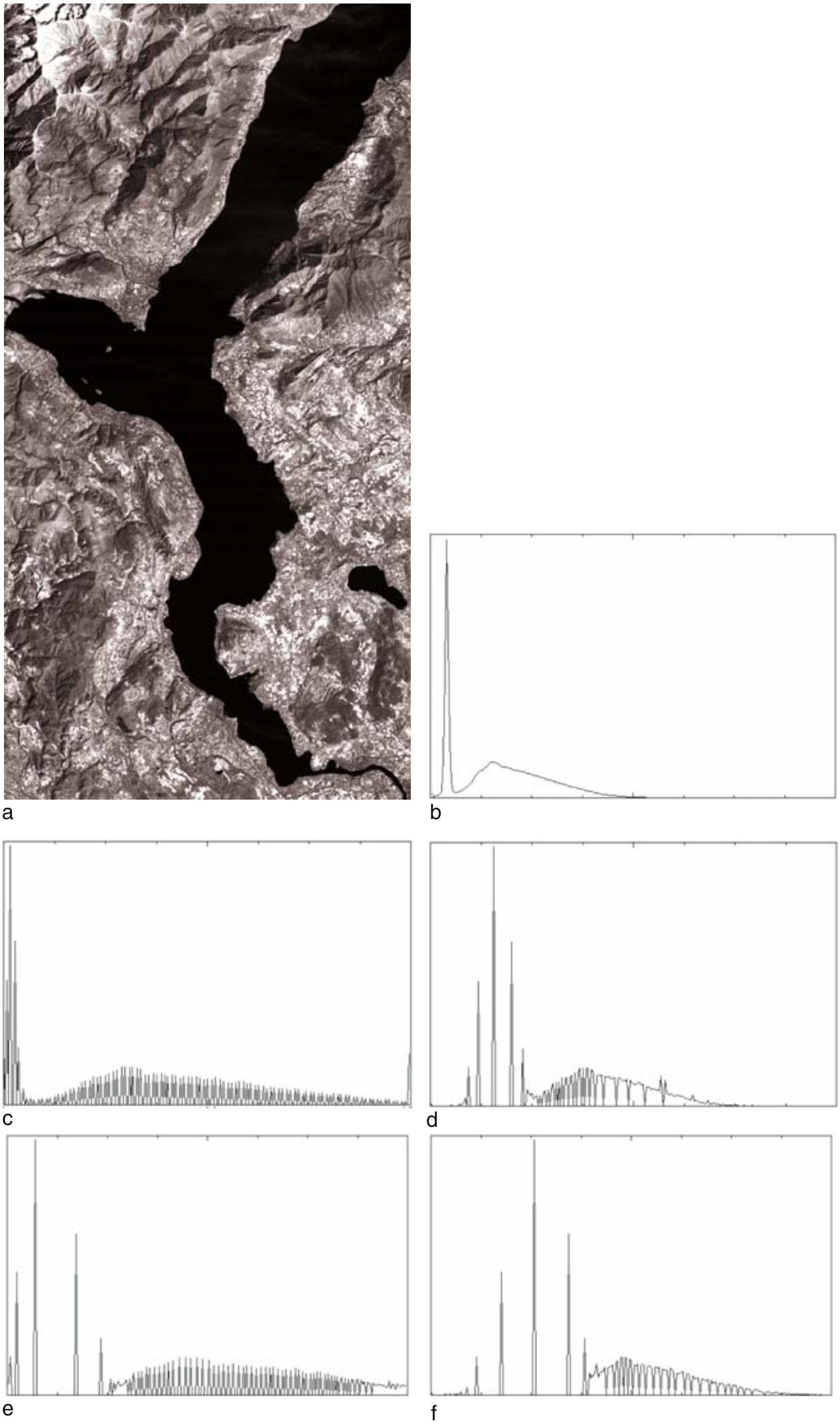


Fig. 8.27 Different types of contrast stretching on an image having a histogram with bi-modal distribution of the data. See Fig 8.26

8.3.1.7 Spatial or Geometric Transformation Techniques

Template Operators

Digital filters are convolution local operators or mobile windows transforming the images in linear and localized mode extracting characteristic elements of interest.

An operator is defined as local if the result of one-pixel application depends on the value of the pixels belonging to its surrounding and eventually by its value. Operators are image transformation methos in which each new pixel's value is calculated in function of the corresponding pixel's value and its position in the original image. A mobile window transformation is invariable by translation. The coefficients of the operation performed on the pixel and its surrounding is represented by a matrix, defined as *convolution nucleus*. The nucleus is centred on each element of the image, and each near pixel is multiplied by the correspondent coefficient; the new pixel's value is calculated on the new surrounding, through the operation defined by the convolution purpose (generally it is a sum).

In convolution operations, the pixel value is replaced with the result of one or more elementary operations performed on the element itself and on the elements of its surrounding; each element of the considered group is multiplied by coefficients which depend on the type of operation to be used (Figs. 8.28, 8.29 and 8.30). These coefficients can also assume negative values. The result will be a new image having the same size as the starting image, where the elements at the borders of the image do not have significant value due to the impossibility in applying the operator completely. To replace each element of the image with its mean the following 3×3 nucleus is applied:

$$\begin{array}{|c|c|c|}\hline 1/9 & 1/9 & 1/9 \\ \hline 1/9 & 1/9 & 1/9 \\ \hline 1/9 & 1/9 & 1/9 \\ \hline\end{array} = \frac{1}{9} \times \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline\end{array}$$

Each pixel covered by the operator is multiplied by the coefficient 1/9; the nine products will be summed and the operation result will define the new grey value of the image element. This operation is repeated by making the convolution nucleus move and centring one by one all the image pixels (*trip mean*). The transformation coefficient's values depend only on the pixels' relative position in the mobile window with respect to the central pixel.

The most used operators are the square: 3×3, 5×5, 7×7, etc. but also rectangular, circular or even more complex shapes are used (Fig. 8.31). In circular operators, only the pixels within a certain distance from the central pixel are used in the computation.

Mobile windows transformations are distinguished into two main classes, in function of the produced image's characteristics:

- *low-pass filters*: transformations providing smoother images than the original ones, i.e. images where the dominant characteristics (low-frequency elements) are highlighted while the details (high-frequency elements) are eliminated;

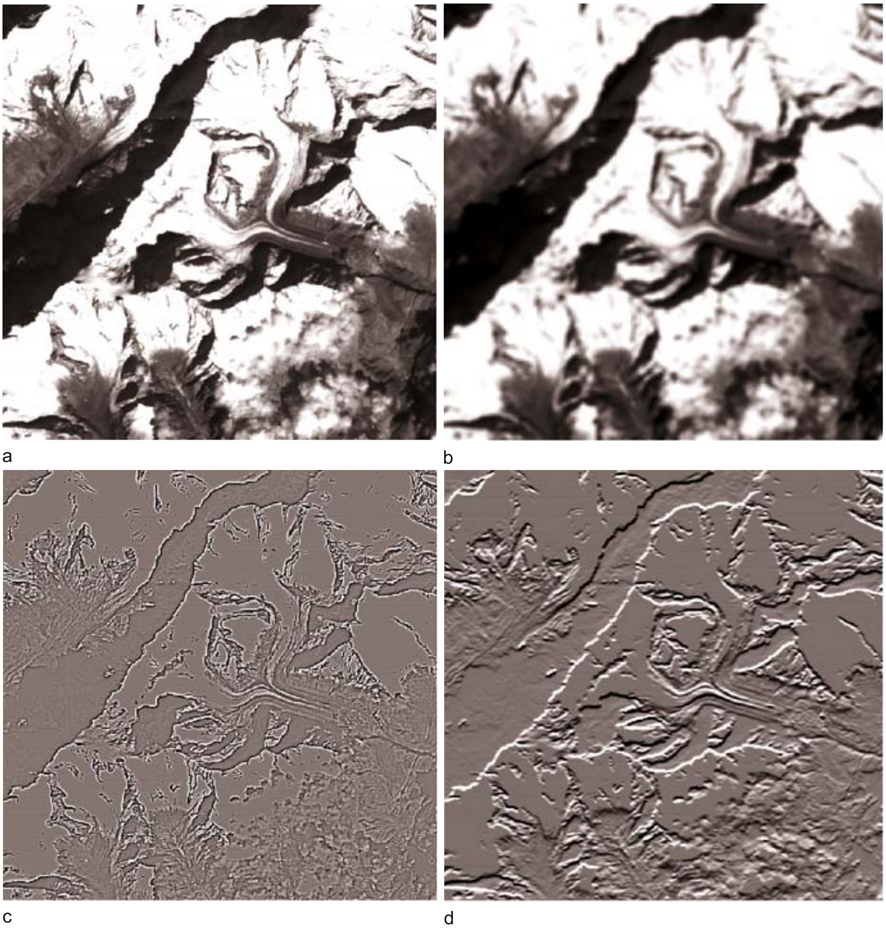


Fig. 8.28 Different types of filters (operators) applied on Landsat TM7, medium infrared band, 2.08–2.35 μm (**a**); (**b**) low-pass, (**c**): high-pass, directional enhancement (+90°)

- *high-pass filters*: transformations enhancing the details, i.e. values variations on small distance (high frequency) eliminating low-frequency phenomena.

While a *low-pass filter* has the effect of reducing the contrast smoothing the borders, a *high-pass filter* enhances elements where the values undergo abrupt changes (Fig. 8.28).

Low-Pass Filter

A simple method to implement a *low-pass filter* is to calculate the mean of the values in a pixel surrounding to be transformed and replacing it with the calculated value.

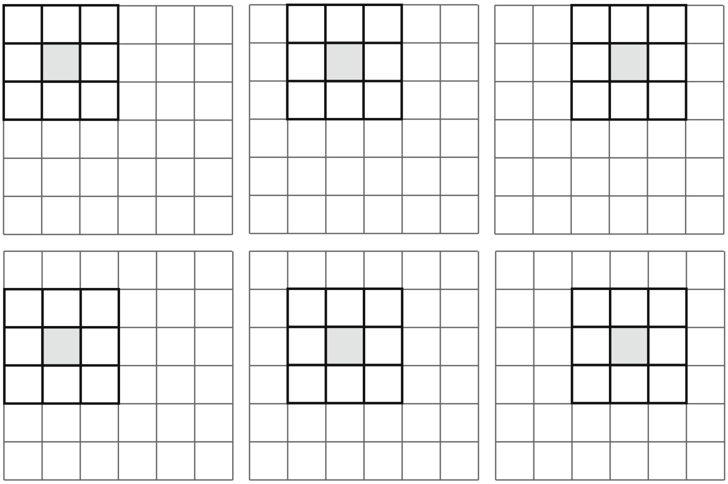


Fig. 8.29 The convolution nucleus moves on all the pixels of the image assigning a new Digital Number based on the characteristic of the filter

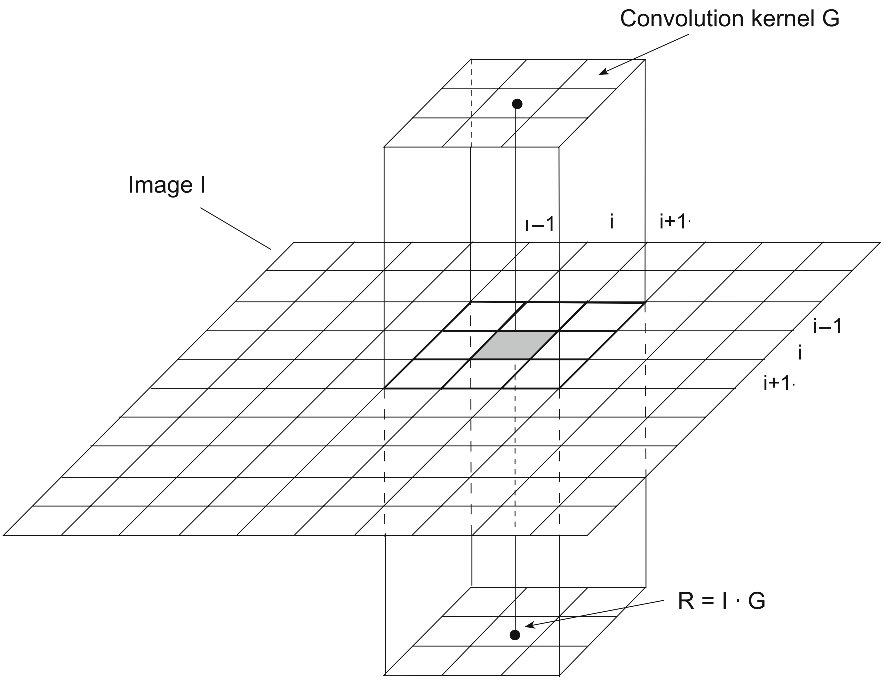


Fig. 8.30 The convolution process is based on the values assigned to each element of the nucleus that computes a new value for each pixel

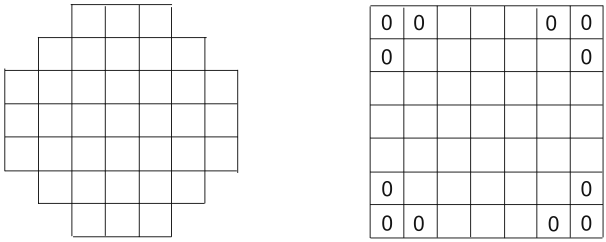


Fig. 8.31 The matrix of a circular filter

It is the most typical case of low-pass filter assigning the mean value of the nearest pixels to the central pixel:

$$\frac{1}{9}\begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline\end{array} \qquad \frac{1}{25}\begin{array}{|c|c|c|c|c|}\hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline\end{array}$$

The operator size affects the smoothing effect's intensity. The *weighted trip mean* is a variant of the previous case. To the pixels of the window are assigned different weights decreasing from the centre towards the borders:

$$\frac{1}{15}\begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 2 & 3 & 2 \\ \hline 1 & 2 & 1 \\ \hline\end{array} \qquad \frac{1}{10}\begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 2 & 1 \\ \hline 1 & 1 & 1 \\ \hline\end{array}$$

The *trip median* is an alternative to the trip mean. The median of a group of values is defined as the central value of the data distribution, ordered by increasing value.

High-Pass Filter

These filters are suitable for the identification of discontinuity, to which high values are assigned, in the images. In smooth areas low values are assigned. The transformation result is very different from the original image, which still allows the identification of the characteristic elements.

Linear elements, faults, rivers, roads, railway lines are well highlighted. The linear elements can be

- borders or boundaries between homogeneous areas, identifiable as imaginary lines;

- one or more pixel wide lines having homogeneous values among them and very different from the nearest.

A simple *high-pass filter* can be implemented by subtracting the result of a low-pass filter application from the original image. One more possibility is in applying a nucleus having a positive coefficient in the central pixel and negative coefficients in the surrounding. For example:

-1	-1	-1
-1	8	-1
-1	-1	-1

-1	-1	-1	-1	-1
-1	-1	-1	-1	-1
-1	-1	24	-1	-1
-1	-1	-1	-1	-1
-1	-1	-1	-1	-1

edge detection: the high-pass operators are suitable to solve the recurrent problem of identification of borders and lines in the image analysis.

The function of these filters is to assign high values to the pixels representing isolated real lines surrounded by homogeneous areas.

In mobile windows, the lines are identified by assigning a higher value along the line, row or central diagonal according to the direction to be analysed.

For this purpose, transition zones between different values in the image are enhanced applying isotropic high-pass filters in the different directions.

These operations, not symmetrical with respect to the pixels to which they are applied, are also called derivative, on the analogy of the differential operation on a continuous function.

A certain border of an image is characterized by null derivative along the border's direction and maximum derivative along the correspondent orthogonal direction (Fig. 8.28).

The following mobile windows are examples for edge detection in horizontal, vertical or diagonal directions:

-1	-1	-1
0	0	0
1	1	1

horizontal

-1	0	1
-1	0	1
-1	0	1

vertical

0	1	1
-1	0	1
-1	-1	0

1	1	0
1	0	-1
0	-1	-1

diagonal

Other mobile windows for lines identification assign negative values obtaining null values in the homogeneous areas. Thus stripes with opposed sign and indication of the direction at the side of the line to be identified are obtained:

-1	-1	-1
2	2	2
-1	-1	-1

E-O

-1	2	-1
-1	2	-1
-1	2	-1

N-S

2	-1	-1
-1	2	-1
-1	-1	2

NO-SE

-1	-1	2
-1	2	-1
2	-1	-1

NE-SO

All types of edges in an image can be identified if the above four examples are applied separately and then combined.

Two or more directions can be identified by the same operator based on Laplace analysis (Laplacian filters) as

-1	-1	-1
-1	8	-1
-1	-1	-1

1	-2	1
-2	4	-2
1	-2	1

obtaining a gradient that considers values modification in the eight nearest pixels.

8.3.1.8 Pan-Sharpening Technique

The technique combining the panchromatic data with the multispectral is called Pan-Sharpening (panchromatic sharpening). It consists in merging the advantages of higher geometric resolution of the panchromatic images with the spectral resolution of multi-band images.

The pan-sharpening operation generates multispectral image with the panchromatic image's geometric resolution.

Transformation techniques can be

- direct and inverse transformation between different colorimetric systems: from RGB (*Red, Green and Blue*), additive synthesis where the generic colour is obtained mixing the primary colours red (R), green (G) and blue (B), to HSI (*Hue, Saturation, Intensity*), a colour representation system using hue, saturation and intensity characteristics;
- colour normalized transformation (*Brovey transformation*);
- colour normalization algorithm (*CN spectral sharpening*);
- *Principal Components Analysis* (PCA).

The panchromatic and multispectral images used in the transformation are overlapped considering the different scale factor, and the multispectral image is first resampled to the panchromatic geometric resolution in order to have the same pixel size.

The direct and inverse transformation between colorimetric RGB-HSI spaces schedules the panchromatic geometric information transfer by replacing the intensity value I within the HSI colorimetric representation derived from the resampled multispectral bands:

$$Multispectral\ image\ RGB \rightarrow HSI + Pan \rightarrow HSPan + I$$

This process works if the original panchromatic DN are normalized between 0 and 1. According to the algorithm, the inverse transformation is performed:

$$HSPan \rightarrow RGB$$

The three resulting new bands have the original panchromatic geometric resolution and the multi-band spectral information if there is correspondence between the original panchromatic spectral ranges and the range resulting from the contribution of the three multispectral bands used in the transformation. The radiometric calibration obtained is always necessary in case of lack of correspondence between the wide spectral range typical of the panchromatic band and the spectral range of the multispectral images.

This correction is performed according to the pixel-to-pixel approach, which gives different radiometric correction for each pixel of the image, only in function of the relations between homologous pixels on the different bands.

Another technique, the Spectral Sharpening transformation technique, can be used with unlimited numbers of images preserving the input original data, radiometric values included; for this reason it is suitable for the fusion of hyperspectral with multispectral images (Plate 8.6).

8.3.2 Qualitative Interpretation of the Images (Photo-interpretation)

Interpretation procedures also remain the same in the time if images are more and more intended as digital and traditional photographs are used with or replaced by multi-temporal, hyperspectral, thermal and radar-band non-analogical sensors.

Although computer-assisted image interpretation and classification are helpful, the activity, experience and skills of the photo-interpreter are fundamental.

The definition of *photo-interpretation* refers to the process by which conclusions coming from the analysis of different elements are drawn.

The *photo-interpretation*, or digital image interpretation, is the extraction of geo-topographic data, territorial and/or environmental information, from photographs, usually aerial photos (photograms), or from aerial or satellite images. The result of the photo-interpretation is the identification and classification of objects.

Photographs are an analogical image produced by the action of the visible light (0.4–0.7 μm) or the visible-near infrared light (0.4–0.9 μm) on a photo-sensitive film (Plate 8.5).

Numerical or digital images are composed of pixels organized by rows and columns in a bi-dimensional matrix. Each pixel, as noted, is described by a number representing its mean radiance and characterized by two plane coordinates x and y.

Photographs and digital images are both *images* although of a different nature, and the prefix '*photo*' in the interpretation is commonly used even though it deals with images.

The photo-interpreter, or image interpreter, by logical procedures detects, identifies, measures and assesses the meaning of the objects, their spatial relations and the models deriving from them.

The digital images acquired by panchromatic or multispectral sensors, as well as the photographic emulsions, have some limits: the interpreter can overcome them based on his own knowledge and by joining his experience to the vision, following a pre-defined observation methodology, finalized to a specific interpretation methodology.

The photo-interpreter has to be able to represent and transfer on maps his perceptions and the meaning of the observed objects, in such a way as to be comprehensible to other users. For this reason the human factor is determinant in the image interpretation (Box 8.7).

Box 8.7 The human factor in the image interpretation

Professional know how	Scholar degree and training	Geographic knowledge
		Historic–social–economical knowledge
		Knowledge in specific domain
Job experience	Interdisciplinary and geographically distributed	
Personal skill	Physical	Visual acuteness
		Stereoscopic vision
	Intellectual	Visual memory Intuition Analytical deductive capacity Analytical inductive capacity Synthesis and extraction capacities
	Character	Observation capacity
	Temperament	Fantasy and curiosity calm e tenancy

Image interpretation is a qualitative process, with the possibility of some metric analysis, opposite to the quantitative process of automatic or semi-automatic classification. The image interpretation needs digital pre-processing of monochromatic or colour images by spatial and spectral transformations which perform respectively georeferencing and contrast enhancement in order for the interpreter to be in the best starting operative conditions.

The image analysis and interpretation process are related to the operator skills, to his experience, and to the ability of merging the single observations in a generalization which is nevertheless subjective.

8.3.2.1 False-Colour Composites

The use of black-and-white images restricts the visualization to the variations along the intensity achromatic axis; black-and-white images do not have either hue or saturation.

By the simultaneous use of the two axes, transforming black-and-white images into a colour image it is possible to observe more details and extract better information.

A colour representation technique is the subdivision according to intensity levels (*density slicing*) in which each grey level or levels interval is represented by a different colour. There are many possibilities in logically assigning colours to grey levels, with two typical cases:

- in photographs, it consists in different layers of emulsion, one for each primary colour (blue, green and red);
- in digital images the colour assignment is based on the use of three tables of transformation (look-up tables).

Each colour can be generated by overlapping the three primary colours RGB in the additive synthesis or by subtraction of the primary colours from the white obtaining yellow, magenta and cyan (YMC) in the subtractive synthesis (see Plate 4.5).

Colour screens use the additive synthesis principle: DNs of multispectral images, stored in three refresh memories, are used to modulate the tension respectively of blue, green and red. When the colour composite uses the three spectral bands RGB of the visible the natural colour image is visualized, otherwise a false colour is represented (Fig. 8.32).

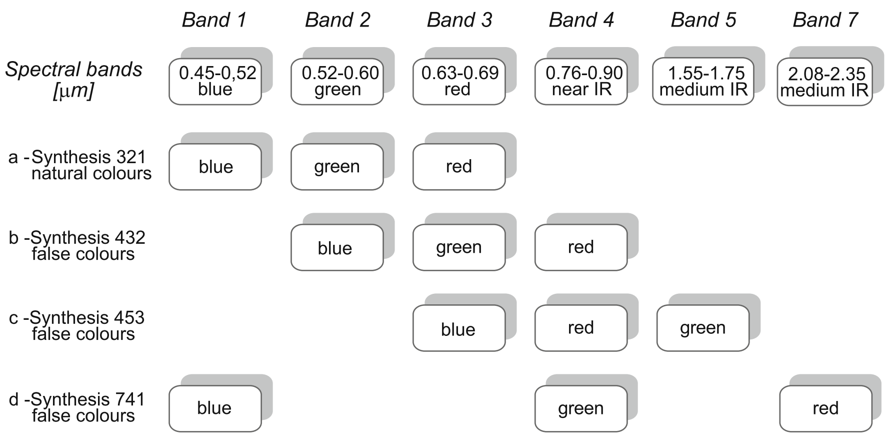


Fig. 8.32 The most used multispectral colour composites of Landsat TM and ETM+; (**a**) RGB:321, (**b**) RGB:432, (**c**) RGB:453, (**d**) RGB:741

The choice of colours to be assigned to a spectral band is arbitrary (Plate 8.7). False-colour composites can be derived from images (i.e. ratios between bands) or non-image type (i.e. altitude maps, population density maps, etc.), on the condition that they are spatially overlapping, generating false-colour hybrid composites.

An example can be represented by the synthesis of two images (obtained by spectral band ratios) and a third made by an original spectral band, which can introduce in the final colour image the information about topographic details removed by the ratios.

8.3.2.2 Visual Processes of the Interpretation

Image interpretation includes at least three visual processes:

- object identification;
- measure of the object on the image or photogram;
- analyses of the observed objects in terms of their meaning according to the research topic.

The third point is complex and requires experience and knowledge of basic information besides synthetical and analytical skills.

- *Identification*: psychological tests measure two kinds of factors: stimulus and feedback.

In photograms or image interpretation stimuli derive from variations of:

– tone;
– structure;
– pattern;
– configuration.

Stimuli determination is easier, while the reply-reaction on the retina, and thus the signal to the brain, is complex (Fig. 8.33). Interpreter sensibility and skills are fundamental.

- *Measure*: photograms' scale factor or digital images' graphic scales are important input for preliminary interpretative–quantitative information. The second necessary element is the knowledge of the real size of the objects. With trigonometric measurement, the size of objects can be derived.
- *Objects analysis* is the interpretation phase; to identify a complex object Φ, it is often necessary to know α, β, γ, recognized as the simplest objects which can, with a high degree of confidence, identify Φ. This phase requires a specific background in the photo-interpreter fields of activity (geology, land use/land cover,

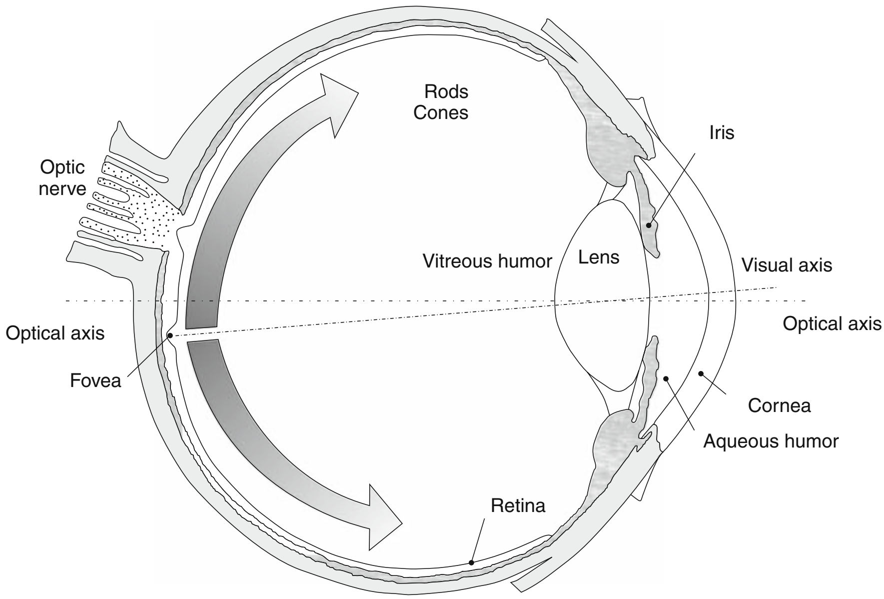


Fig. 8.33 The crystalline lens in the eye acts as a lens tube of a photo camera, while the pigmented iris regulates the incoming light intensity on the retina controlling the variable aperture of the pupil (the gain in instrumental terms). The pupil has the same role as the lens stop (diaphragm) in the photo camera; cones and rods are the photo-detectors, namely light sensor cells, equivalent to photographic film, optical-mechanics and digital detectors

geomorphology, planning, etc.) combined with a significant experience; in fact many factors may induce interpretation mistakes. The graphics shown in Fig. 8.34 are explanatory examples.

8.3.2.3 Elements in Image Interpretation

The *basic elements* for photograms and digital image interpretation are as follows, considering tone and colour as different aspects of the same element (Fig. 8.35).

- *tone and colour*, basic element;
- *size and shape*, interpretative elements of primary importance, geometrically defining tone and colour;
- *height and shade*, geometric interpretative elements related to the previous two, but less important;
- *structure and texture* (or pattern), primary importance interpretative elements, spatially defining tone and colour.

The first one, *tone and colour*, is the most important as it is not possible to observe and analyse the other elements without a good contrast in the image to be interpreted. This element is related to the object's spectral response and to how

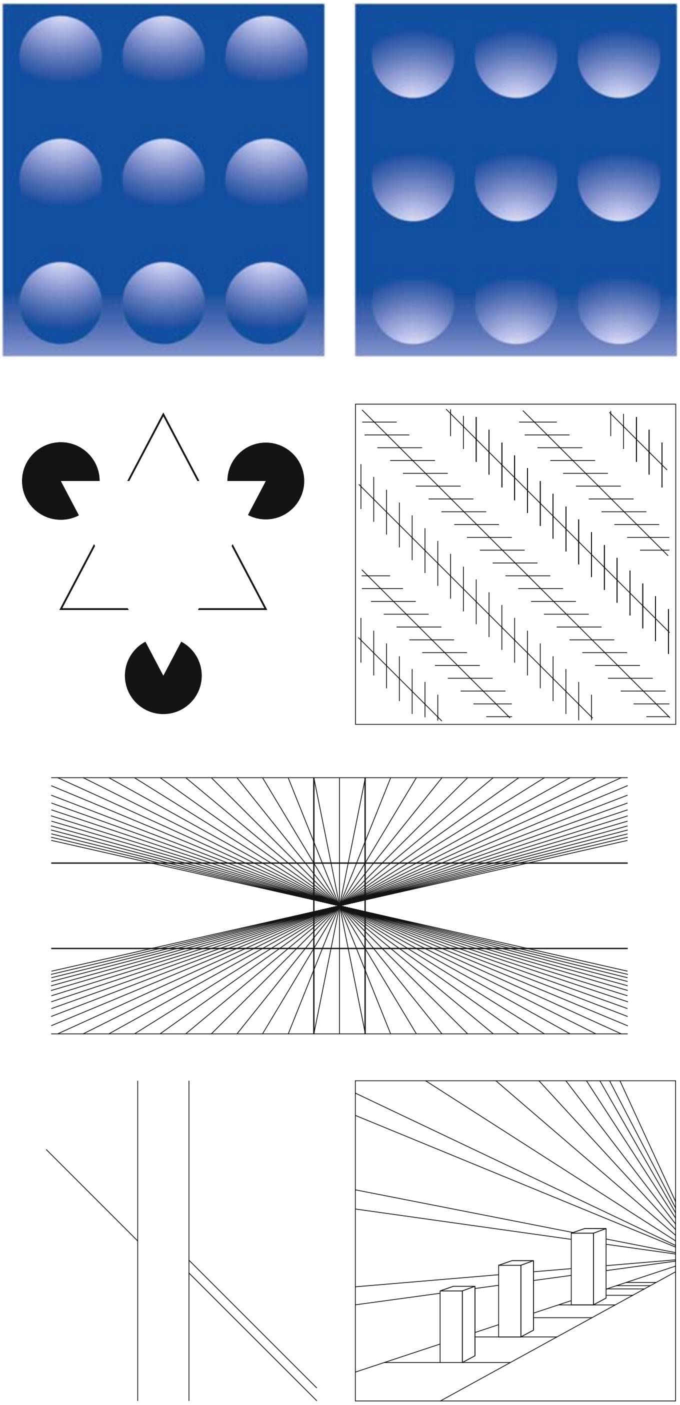


Fig. 8.34 (continued)

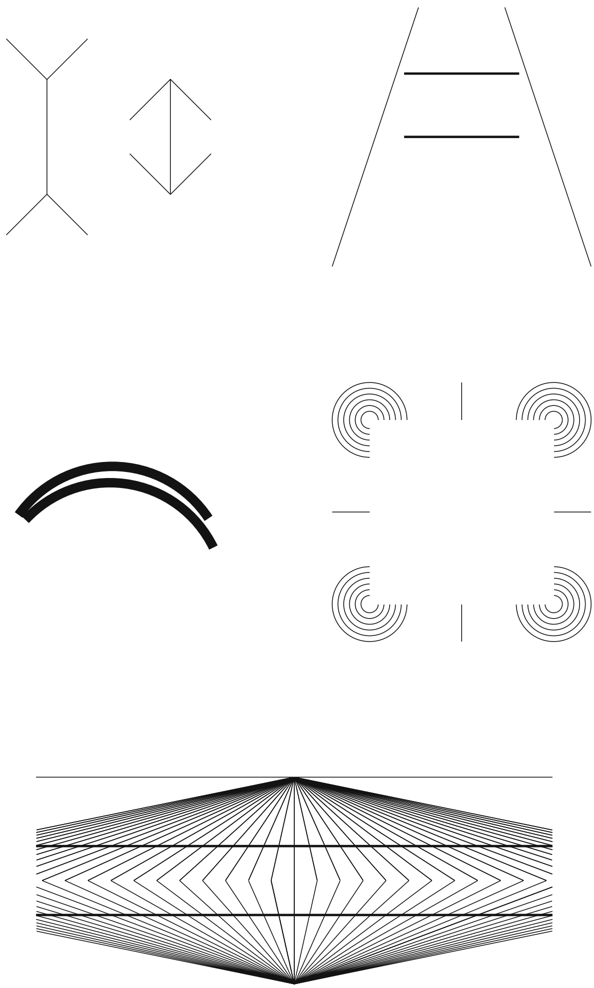


Fig. 8.34 The view could be wrong; only with an accurate observation is it possible to correctly read the graphical examples; the same objects look different if associated with other graphical signs (shadow, lines, perspective), or not-drawn figures appear

this is represented in the image. The other elements can be divided into two groups: *geometric* and *spatial*. The most important are the first two geometrical, *size* and *shape*, and both the spatial, *structure* and *texture*. *Height* and *shade* are secondary in the scale of priority adding other geometric information to the knowledge coming from the elements previously considered.

Geometric and spatial elements depend on the images' resolution characteristics, i.e. on the resolving power of images at a given scale.

The other complementary elements to better characterize the interpretation are

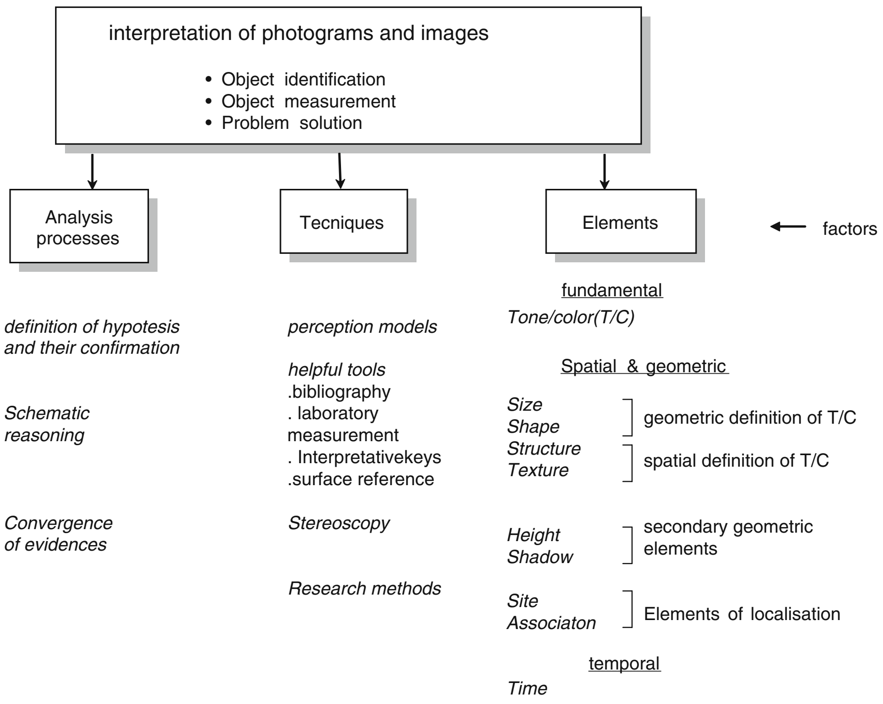


Fig. 8.35 The qualitative analysis and interpretation of photograms and imagery

- *location and association*, elements related to the object absolute and relative (with respect to the other objects') position;
- *time*, related to the time when the object is observed, or as analysis of its variability.

These elements are perceptive abstractions; they are very much dependent on the photo-interpreter sensibility and experience. They describe the objects not in terms of how they appear, but in terms of the relations they have with the surrounding environment.

Tone and Colour

The basic element for the image interpretation is the *tone*. The tone contrast distinguishes the objects in a scene.

In colour images, the tone is represented as different *hue*, *intensity* and *saturation* combinations. The colour is a basic characteristic of natural elements, while it can be less important for artificial elements.

In black-and-white images, the tone is represented as brightness that increasing, through grey levels, can go from the black to the white. In a digital image, the tone is subdivided into grey levels whose highest number depends on the radiometric res-

olution measured in bits. The tone corresponds to the intensity level in a monochromatic image. In black-and-white images (B/W), the interpreter associates grey tones to different objects, knowing the external factors that can have an influence: season, illumination, latitude, etc. All the objects can assume different grey tones in function of the Sun angle, the surface morphology and the wavelength interval.

The keys for image interpretation based on tone and colour are in the following elements:

- how each object reflects or emits the electromagnetic energy expressing it in different tones;
- how the tones vary according to the wavelength;
- which relation exists among the objects' spectral properties and the spectral sensibility of photographic films or digital sensors.

Colour perception is an important factor for image interpretation. Objects reflect, emit and transfer different quantities of energy at different wavelengths (λ). These differences are collected as tone, colour or density variations on images.

The human eye is sensitive only to the visible spectrum, whose wavelengths are perceived as colours that can be considered as combinations in hue, intensity and saturation of the three primary colours' RGB synthesis. Colour images' interpretation always refers to RGB combinations, at any wavelength. Objects' interpretation out of the visible spectrum is not intuitive. The photo-interpreter has to know the objects' spectral behaviour in the different ranges of the electromagnetic spectrum and how these behaviours are influenced by the environmental factors.

The use of false-colour images is commonly used, increasing the interpretability in the case where one or more infrared bands are represented in one of the main colours.

In a thermal infrared image, in the emitted domain, the occurring tone variations are related to differences in the surface temperature; by the use of look-up tables these tones can be transformed from greytones into colours.

Colour tones provide more information about the objects and the observed scenes than the only greytones.

The human eye can distinguish a higher number of colour combinations (hue, intensity and saturation) than grey levels. Black-and-white images in fact restrict the possibility of observations to the only intensity scale, as they do not have hue and saturation (Fig. 8.36). Anyway to distinguish an object good contrast between the object and the background is necessary.

Black-and-white images do not help the perception of the interpreter, who remembers the objects in a colour reality. Near infrared images ease interpretation and discrimination of some elements, enhancing the contrast with the surrounding objects. The use of selective filters improves the quality and contrast of imagery and the interpretation: blue filters on photo camera lens acquired with a colour film smoothes the atmospheric scattering reducing the out-of-focus effect produced by the atmosphere.

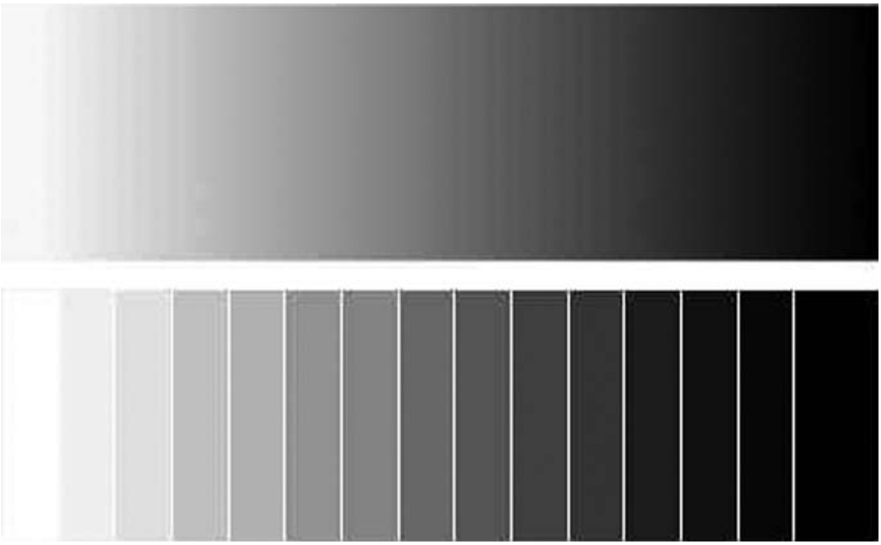


Fig. 8.36 Continuous grey scale (above) and segmented grey scale (below). The separation of the continuous grey tones is very difficult for a non-expert eye

Size or Dimension

The size refers to the area occupied by an object, but can also refer to its linear dimension. It is, with the shape, a higher order geometric element of interpretation. *Size* with *shape*, *height* and *shade*, forms the basic group of image-interpretation criteria for the characterization of the geometry of objects.

The identification of the objects' size helps the interpreter to fix a reference scale and, in defining in non-rigorous terms the size of other objects (Fig. 8.37). Their measure can help the interpreter to exclude some uncertainty in the identification

Fig. 8.37 The size of different objects in a scene contributes to their interpretation. In the centre: circular agricultural field with rotating wing irrigation system

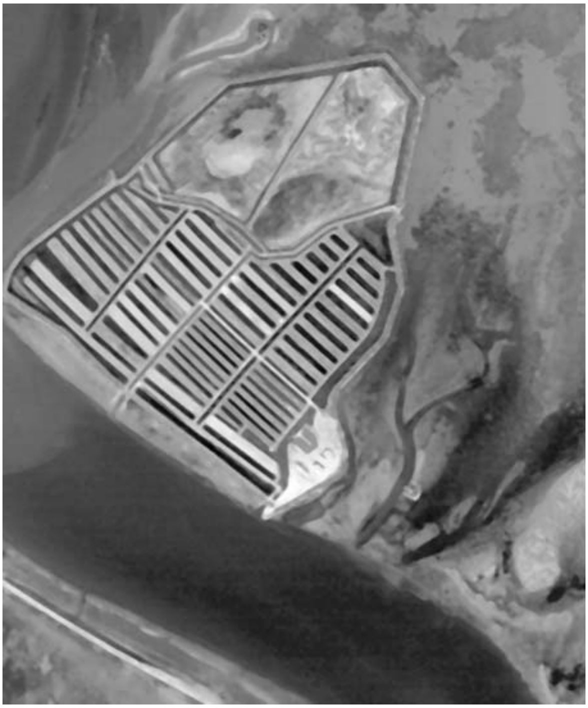


Fig. 8.38 Different size of cultivated field in a lagoon area

(Fig. 8.38). A 3D view of stereo-pairs may help in defining the height of the objects and carrying out parallax measures useful for indirect geometric characterization (buildings, tree size). The volume that can be obtained from a poplar wood depends on the trees' size (height and diameter), the trees' density (trees per hectare) and the plantation size.

Shape

The shape is defined by geometric borders and provides information not only on the identity of the object but also on its function and meaning. For example, the shape can be related to the factors that determine it: dune morphology and the side colonized by the vegetation can provide useful elements to define the prevailing wind direction. The round, long, linear, punctual, etc. shape is typical of objects and allows their identification as the association of shape to objects is often univocal. Some examples can be gardens, golf courses, a leisure area or football field, road network or motorway connection, a rice field highlighted by its borders, water meadow and secular olive trees.

Fig. 8.39 The shape of the buildings in an urban area defines a structure

However this relation cannot be valid if the object is observed from two different altitudes so that an unskilled interpreter cannot recognize objects usually seen from a terrestrial perspective. Vice versa the aerial view is often determinant: such as a motorway entrance, easy to recognize for a photo-interpreter, or a labyrinth for the driver.

Once the object perspective is defined, important information about its structure, composition and function can be retrieved. The illusion of the higher depth of the objects in stereoscopic vision can help to complete the problem solution and increase the photo-interpreter skills (Figs. 8.39, 8.40 and 8.41).

Height

The *height*, or its opposite *depth*, like the shade, is a secondary importance interpretation geometric element. Skilled photo-interpreters can calculate the objects' height or depth from the stereoscopic vision of images, without using parallax instruments needed for an accurate measure. As mentioned above, comparing known objects' heights can help in the identification and the measurement. This way elevation and depth can be assessed (Figs. 8.42 and 8.43).

Fig. 8.40 The shape of an island in the texture of canals in the Grado Lagoon, Adriatic Sea, Italy

Shadow

This secondary interpretative element not only can be useful to identify objects but can also hide details for the interpretation. The objects' size and shape can be derived by observing their shadow, especially in the case of lack of stereoscopic vision. Its observation is discriminative if the objects are very small and if there is a low contrast with the background, but it is not detectable if it is projected on dark surfaces or water bodies. The shadow changes according to the time and orientation of the acquisition, to the season, hence it depends on the azimuth and on the Sun height, and by planning these factors it is possible to reduce or increase its relevance for the interpretation. Usually the aerophotogrammetric acquisitions are carried out in the spring–summer period during the maximum Sun energy around midday, to reduce as much as possible shadow effects. Visible and reflected infrared satellite images are generally collected in the interval 9.30–10.30am, the time of passage over the equator. This time has been chosen as related to the presence of less clouds and mist and undesired reflection effects, phenomena limiting the acquisition capacity of passive optical instruments, nevertheless generating elongated shadows, with stronger effects in winter months (Fig. 8.44).

Fig. 8.41 The shape of a highway interchange

NOAA satellites have acquisition cycles at 7 am and 2 pm. These parameters are constant by the seasons and not changeable, thus they have to be considered during the interpretation. Moreover it is important to consider if the shadow of the object is on a flat area or on ascending or descending slope: its projection and measure are a function of the slope (Fig. 8.45).

The active SAR lateral vision, fixed at 23° on ERS 1–2 satellites and variable between 20° and 49° on Radarsat, can help in identification (electric pool line, poplar wood border, water bodies) or not (shadow in mountain zones hiding much information).

Then there is the problem of the presence of cloud shadow on the images (Fig. 8.46). If aerophotogrammetric flights are under the clouds darker tones on the shadowed areas with sharp or smoothed tone boundaries are generated. The shadowed areas can be processed with specific masking (Fig. 8.47).

Structure or Pattern

Structure is, with the texture, a geometric element of primary importance as it recognizes the objects according to their spatial position. The structure, or *pattern*, in an

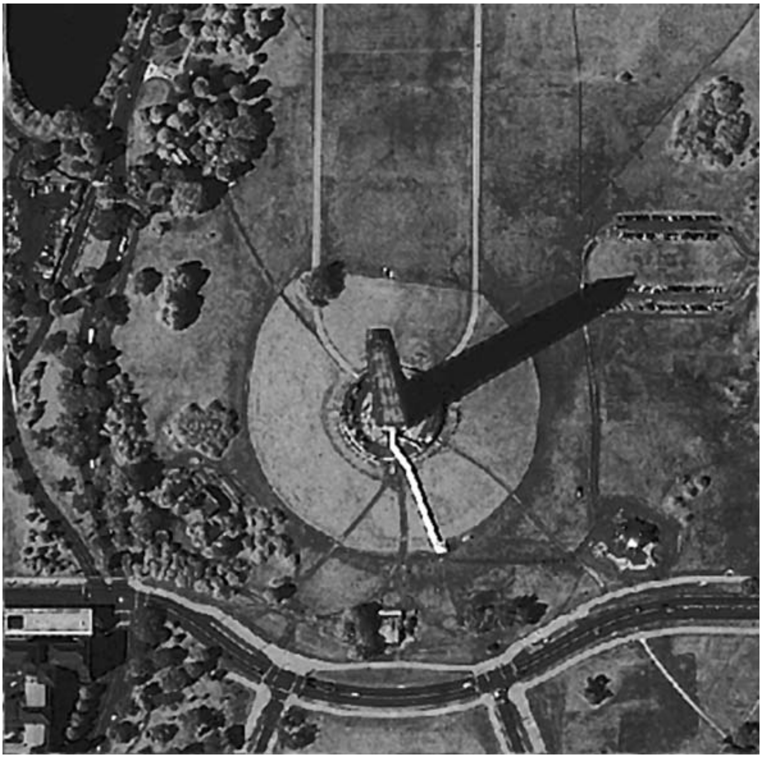


Fig. 8.42 The height of the buildings and of the trees could be indirectly derived from the obelisk shadow

image is given by the spatial distribution of the tone repetition of groups of objects, both natural and human made, often too small or insignificant to be distinguished as single entities. The objects are set together in order to form a pattern that aids recognition by the interpreter (Fig. 8.48).

The hydrographical network patterns are typical, related to different lithology, geological structure or soil type situations.

The trees' size, for example, is interpreted based on the apparent structure which changes according to the images' scale. On a large scale, in a wood each tree can be identified, but not its leaves, which nevertheless contribute to define the tree canopy structure; on a small scale the canopies of the same wood contribute to form its structure. To an appropriate scale, the spatial distribution and the frequency of leaves', branches' and stems' tone borders provide information on the type of tree species as they define its pattern: typical is the difference between a beach and an olive tree structure. The parking around a shopping centre has a defined pattern, often in correspondence of an important road.

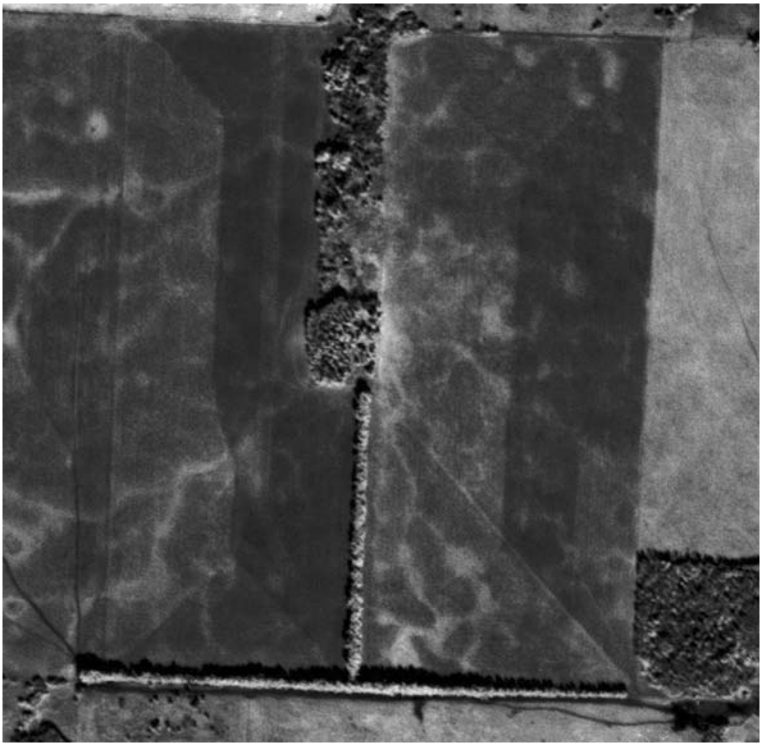


Fig. 8.43 The changing height of the horizontal line of trees is measurable from the shadow projection on the soil

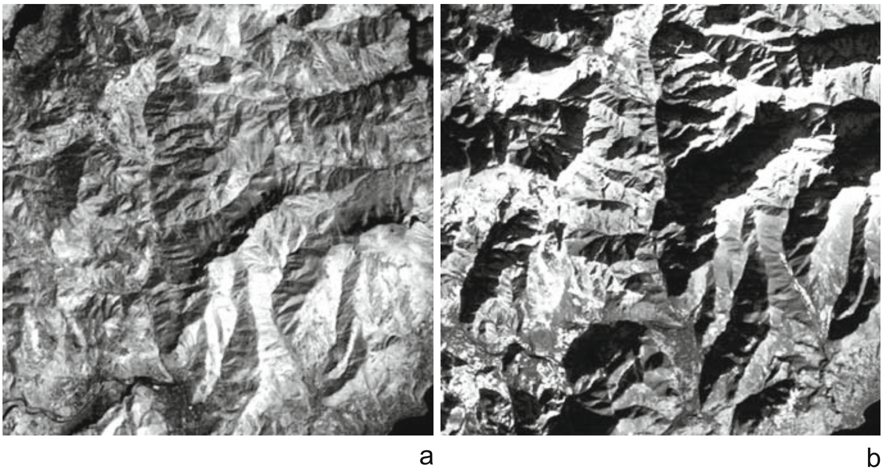


Fig. 8.44 The same mountain scene in summer (**a**) and in winter (**b**); the shadows are more elongated in winter, causing an important limitation in the interpretation phase

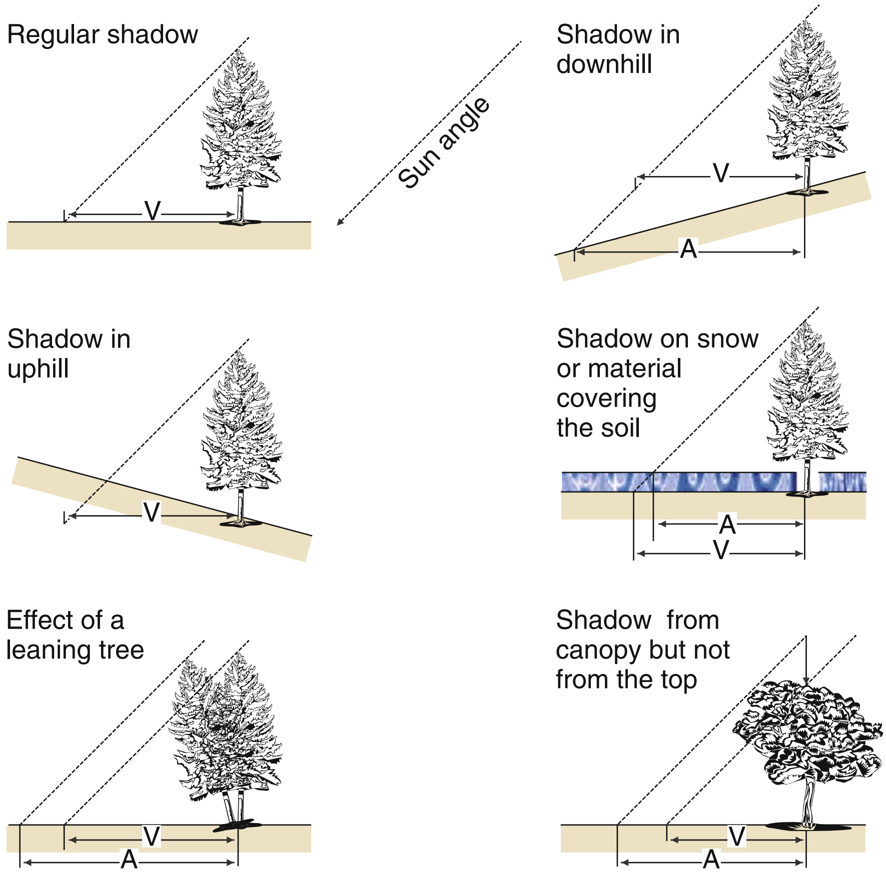


Fig. 8.45 Different possible situation that can occur in the interpretation phase; the length of the apparent shadow (A) is different with respect to the true shadow length

Texture

The *texture* can be defined as a micro-change in an image tone distribution, i.e. the quality discriminating between two objects having little or null tone differences (Fig. 8.49).

The morphological characteristics play an important role in texture. Elements such as irregularity or roughness, smoothness, exposition and orientation, superficial stratification, soils and rocks permeability, types and density of vegetation generate characteristic lights and shadows in the images. An example is the smooth or rough aspect of an observed scene: two objects with medium grey levels in a panchromatic image, as a water body and a wood, are distinguished for their texture difference (Fig. 8.50).

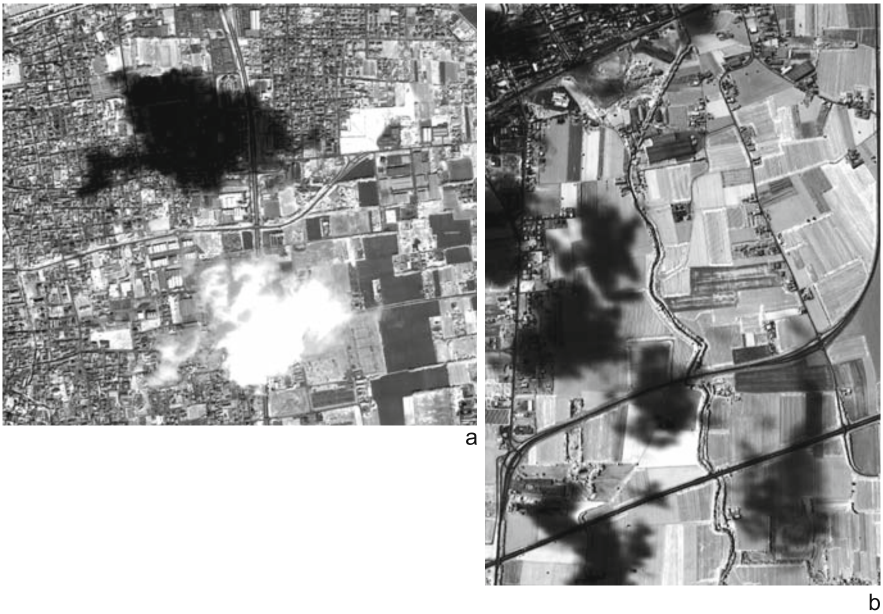


Fig. 8.46 (**a**) Cloud and its shadow in QuickBird panchromatic image with 0.61 m nominal geometric resolution; (**b**) cloud shadow recorded from the hyperspectral MIVIS sensor, nominal geometric resolution of 2 m

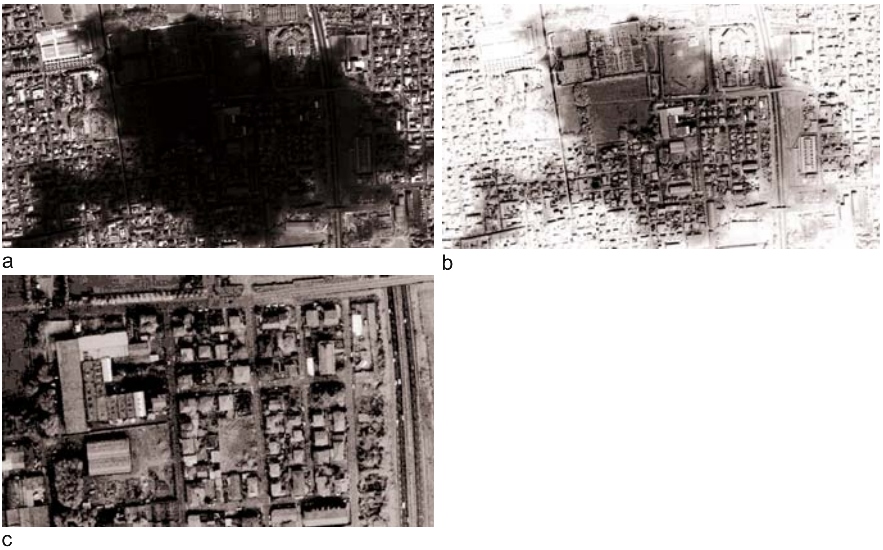


Fig. 8.47 Large shadowed area that modifies the radiometric behaviour of the pixels. With an appropriate contrast enhancement, the shadowed area can be interpreted bringing to saturation the all round area; (**a**) shadowed area; (**b**) shadowed area with Gaussian contrast stretching; (**c**) zoom in the shadowed area with high contrast useful for the interpretation

Fig. 8.48 Structure of a railway station with wagon lines on the railroad

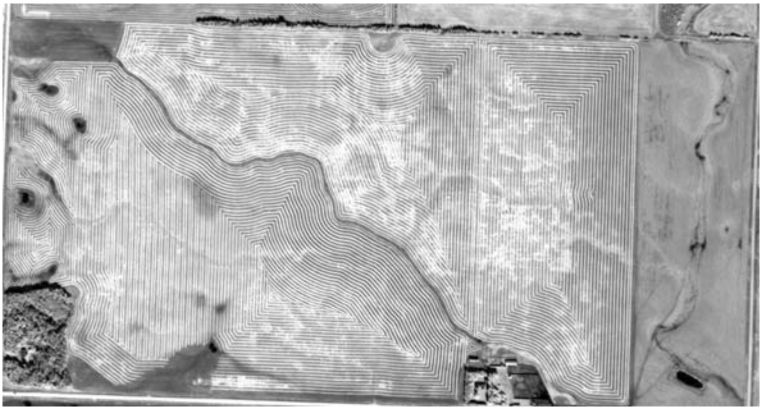


Fig. 8.49 Tonal repetition and light/shadow alternation define the texture of the ploughed field in the image

The photo-interpreter can appreciate the differences among apparently similar objects in a scene if he/she can understand the relationship between the texture and its origin in the environment. The texture generated by human activities is characterized by linear and/or angular shapes. The different relations between objects and the environment produce characteristic textures: herbaceous and arboreous cultivations, ploughed fields, irrigation systems, as well as a railway more linear and with fewer intersections than a road. Hydrographical networks can differ in function of the lithology, the slope presence or the absence of biomass. Structure and texture analysis are facilitated by the synoptic properties of airborne or satellite remotely sensed images.

Fig. 8.50 The different texture of the natural landscape allows resolution of classes with little tonal differences

Site

The site indicates both the geographic localization and the position of an object with respect to the others, as the environment of an element contributes to its identification.

The site, together with the other complementary elements *association* and *time*, supports the relations between objects, phenomena and environment; these elements describe the objects not as they appear but for the relations they have with the surrounding environment.

Knowing the objects' geographic position the number of possible interpretations decreases and more information, not necessarily derivable from the images, can contribute to the identification. Aspect, geology, soil type, vegetation, latitude, macro- and micro-climatic conditions are important site factors. Some tree species grow well in marshes and others in dry areas; in medium-high mountains only some vegetation typologies can be expected to live.

Association

Association defines the spatial relation between objects and phenomena. The *association* is an abstraction of the objects' proximity and connection and of phenomena in a certain environment. Some objects are commonly associated with others so that often one confirms the presence of the other. A dense network of canals indicates a link with irrigation practice, an artificial lake can be associated to a hydroelectric power station, the urban structure changes in different cultural areas, in Mediterranean environment shrub vegetation can be associated to the Mediterranean maquis.

Time

This element of image interpretation relates the multi-temporal relationship of objects and phenomena. Some objects can be better identified if their temporal changes are considered, i.e. the vegetation phenological cycles (Fig. 8.51). Changes occurring in a few hours can be associated to meteorological events or disasters, like landslides, earthquakes, fires. Short-period (weeks or months) changes include variations due to human interventions or agronomical cycles. One or more year medium-term changes can include buildings and infrastructure modification, or slow evolution of natural phenomena.

The time is also related to the moment in which an observation is carried out: some objects or ground situations can be better seen in winter when the biomass does not constitute an obstacle. The element *time* has become relevant since multi-temporal acquisitions have been systematically available. The radar permanent scatterers technique is a typical example.

8.3.2.4 Resolution

Sensor physical parameters, i.e. geometric, spectral, radiometric and temporal resolutions, define the interpretation limits of the images. Being the instruments used for the acquisition equal, the resolution depends on the altitude of acquisition (see Chapter 6).

In analogical stereoscopic photogrammetric acquisitions, the geometric resolution is a function of the resolving power of the photographic system, of the flight plan characteristics, of altitude and camera focal length, which define the scale of the photogram, and environmental conditions. The same areas acquired after some days can have different resolutions due to changed acquisition conditions. The resolution varies also in the same photogram according to whether the central part or the border of the image, a mountain top or its valley bottom, are observed. The resolving power, defined as the number of parallel lines that can be separated in 1 mm (lines/mm), depends, besides on the photographic emulsion characteristics, on the reference objects' contrast. A low-contrast system can resolve 30 lines/mm, while a high contrast can separate 60 lines/mm; the ground resolution distance (D_r), measured in cm, hence derives from the combination of the images' scale and the resolving power:

$$D_r = \frac{\text{image scale}}{\text{resolving power}} \cdot 1000 \qquad (8.9)]$$

In images acquired at 1:1500 scale, in low contrast conditions (30 lines/mm), D_r corresponds to 5 cm and, in high contrast conditions, to 2.5 cm.

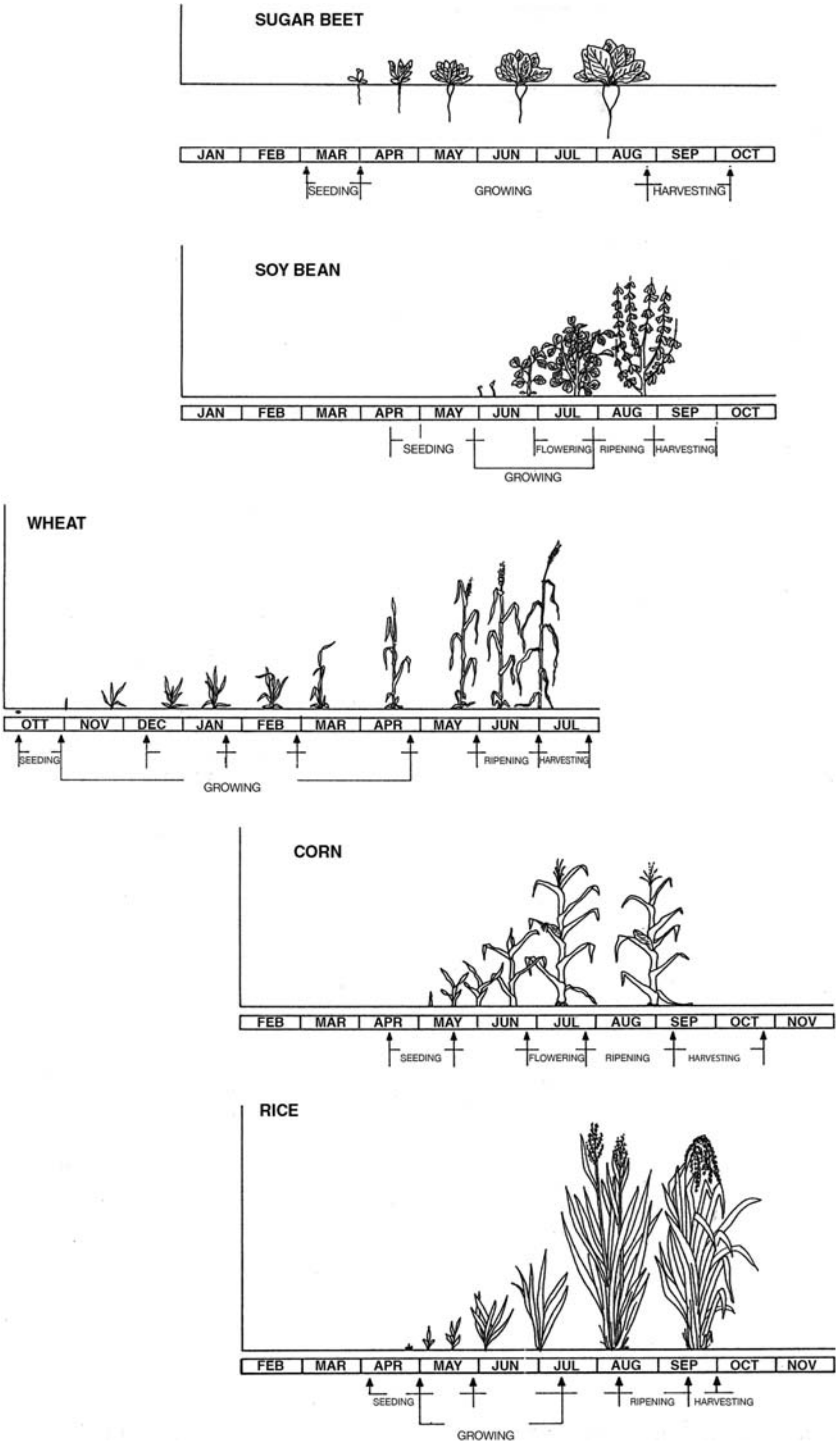


Fig. 8.51 Phenological cycles of some crops compared for a multi-temporal analysis

The relation between the observed object and its surrounding is very important. A white object on a black background is much better resolved than a bush in a grass field.

These observations are valid also for digital images, where the signal corresponds to the ground minimum and unitary surface, the pixel. Defined as ground resolution, or Instantaneous Field of View (IFOV), the pixel varies, according to the acquisition system's characteristics, from some decimetres to kilometres.

The pixel size affects the detail interpretation in imagery.

Moreover metrical analysis can be carried out on the images to determine different parameters; the most common are shown in Box 8.8.

Box 8.8 Metric analysis of images interpretation

Parameter			
a'	Photo side	d_i	Distance on the image
a	Field of View (FOV)	DS_i	Denominator of the image scale
f	Camera focal distance	DS_c	Denominator of the map scale
h_g	Relative altitude: distance acquisition system – ground level	d_c	Distance on the map
h_o	Absolute altitude: distance of acquisition system – average sea level	Δh_1	Known altitude
h	Terrain elevation, geoidic height	Δh_2	Altitude to be determined
S_i	Image scale	l_1	Shadow length of Δh_1 on the image
s	Known size of an object on the ground	l_2	Shadow length of Δh_2 on the image
s'	Size of the same object on the image		

Parameter to be determined	Formula
Relative flight height (h_g)	$h_g = h_o - h$
Field of View (a) at a relative flight height (h_g)	$a = \frac{h_g}{f} \cdot a'$
Absolute flight height (h_g) to obtain the Field of View (a)	$h_g = \frac{a}{a'} \cdot f$

Parameter to be determined	Formula
Image scale factor: - focal distance (f) and relative height (h_g) ratio - comparison with a known object size - comparison image – topographic map	$S_i = \frac{f}{h_g} S_i = \frac{s'}{s} S_i = \frac{d_i}{d_c} \cdot DS_c$
Object height on an image shadow comparison - other methods: radial displacement Shadow factor (tangent – height) parallax differences	$\Delta h_2 = \frac{l_2}{l_1} \cdot \Delta h_1$
Velocity of an object	Computation of the distance knowing the time (ΔT) between two sequential stereoscopic images

8.3.2.5 Vegetation Spectral Behaviour

The vegetation has a different behaviour at each spectral band. The reflection of the incident electromagnetic (EM) radiation for each band contributes to the definition of typical spectral responses. The curve of reflectance can vary in function of many factors such as vegetation type (macro- and micro-phyta, broad-leaved and coniferous, etc.), density, phenological phase, phytosanitary condition, moisture.

In particular its spectral behaviour depends, in the visible and medium infrared bands, respectively on the leaf pigments' content and type, on the leaf structure and on the water content (Fig. 8.52).

- *Leaf pigments*: the EM energy reflected in the visible is related to the presence of leaf pigments like chlorophyll, xanthophyll and carotene. In particular the chlorophyll determines the absorption in the blue and in the red and a reflection of the incident radiation, about 8–15%, in the green. The epidermis and the pigments are transparent to the infrared, allowing this radiation to reach the parenchyma. The green and yellow radiation components (not the red one) are reflected by the mesophyll, giving the typical green colour in the visible.
- *Leaf structure* is responsible of the spectral behaviour in the near infrared bands, in the range 0.70–1.35 μm inducing very high reflection of the incident energy, of the order of 30–70%. The residual energy is mainly transmitted, while the absorbed energy is much reduced. In remote sensing, this phenomenon is at the base of the most relevant analysis about vegetation. In fact, as the mesophyll development is linked to the specific variety (cultivar), to the plant health conditions and phenological phase, the study of the reflectance in the near infrared is a powerful investigation instrument to distinguish different species and to analyse the evolution phases of a homogeneous vegetation community.
- *Water content*: in the medium infrared domain, 1.35–2.70 μm, vegetation spectral properties depend on the leaf water content. In water stress conditions, there is

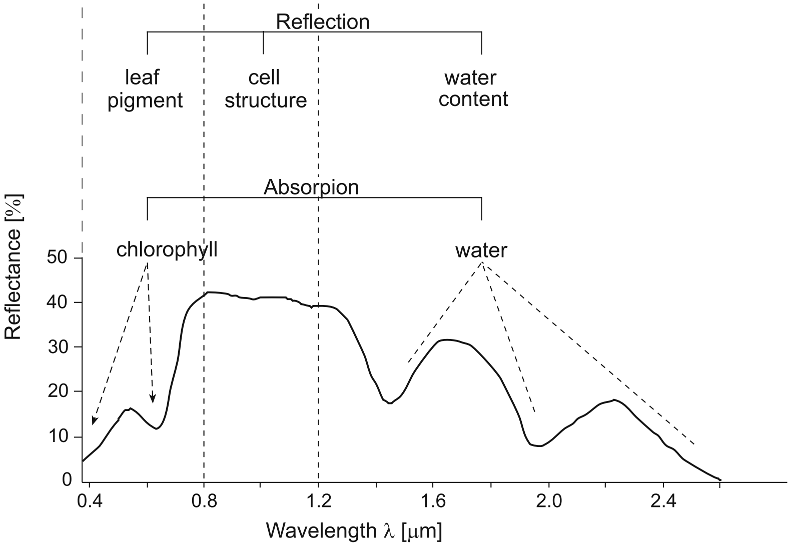


Fig. 8.52 Relationship between typical reflectance curve, but not univocal, of the vegetation in different electromagnetic spectrum regions and leaf pigments (in the visible), leaf structure (in the near infrared, water content (medium infrared)

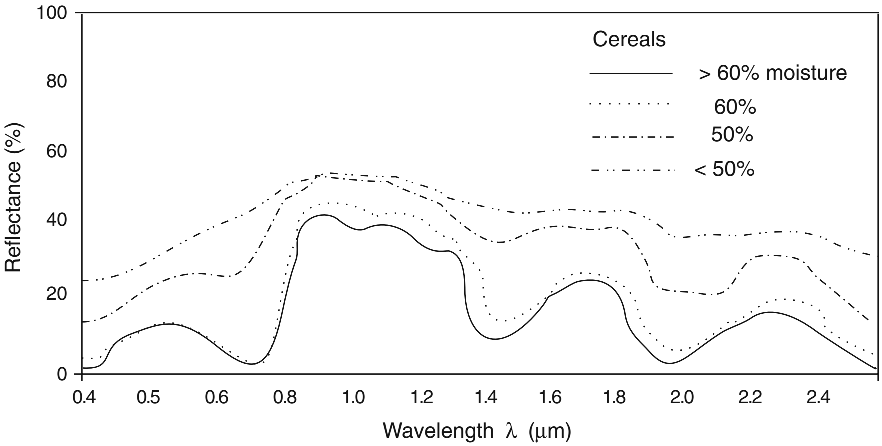


Fig. 8.53 Spectral response of the vegetation with different water content

an increase in the reflected radiation values, stronger in some absorption bands (Figs. 8.53 and 8.54).

Figure 8.55 shows a schematic representation of the vegetation electromagnetic radiation reflection, transmission and absorption mechanisms. Also the lower layers, although in small part, contribute to increase in quantitative terms the reflected

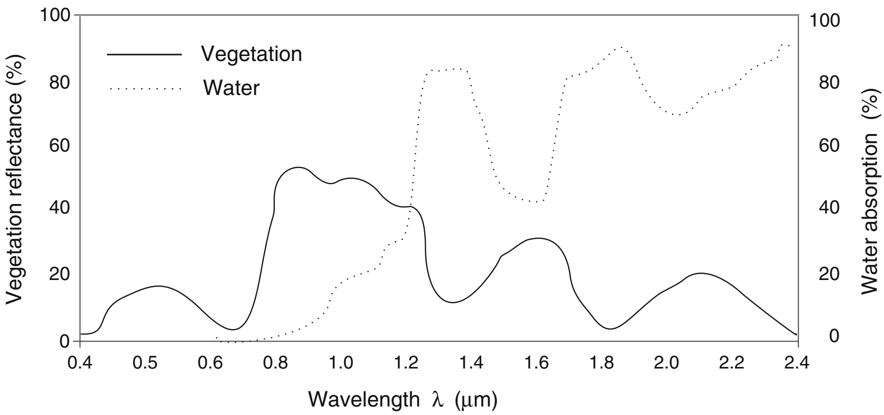


Fig. 8.54 Typical spectral response of the vegetation compared with the water absorption curve: in presence of a higher absorption, the energy that reach the leaves decreases and consequently the reflectance decreases

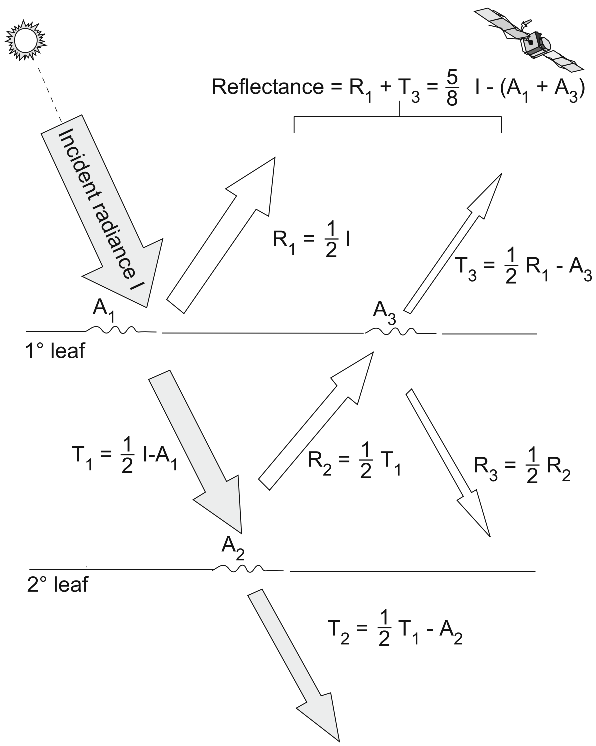


Fig. 8.55 Quantification of the reflection, transmission and absorption mechanisms of the incident radiant energy on the multi-stratified leaf biomass. I: Incident Energy; T: Transmitted energy; R: Reflected energy; A: Absorbed energy

energy. The spectral behaviour of different species is different according both to plants' and leaves' structure.

The quantitative relationships between reflected energy and phenology can be defined through algorithms based on the ratio among spectral bands. Among these there are the Vegetation Indices and the use of statistical units like the Tasseled Cap transformation (Kauth & Thomas, 1976) and the Principal Components analysis (Richards & Xiuping, 1998).

Vegetation can act as a superficial and under-the-surface indicator of environmental modification. Moreover it is one of the easier classes to distinguish by remote sensing, due to its typical spectral response.

To better understand the differences in vegetation spectral behaviour, the typical example of the needle leaved and broad-leaved is often presented: as they have different leaf structure they also show remarkable differences in the incident energy reflection (Fig. 8.56 and Plate 8.8).

The reference panel in the two photographs (Plate 8.8) indicates how the colour changes due to the false-colour film effect compared to the colour films. In broad-leaved the brighter magenta colour with respect to the needle leaved indicates a higher reflectance of light radiation in the photographic near infrared (0.7–0.9 μm), with percentages around 45–50%, similar transmittance (45–50%) and absorption lower than 5%.

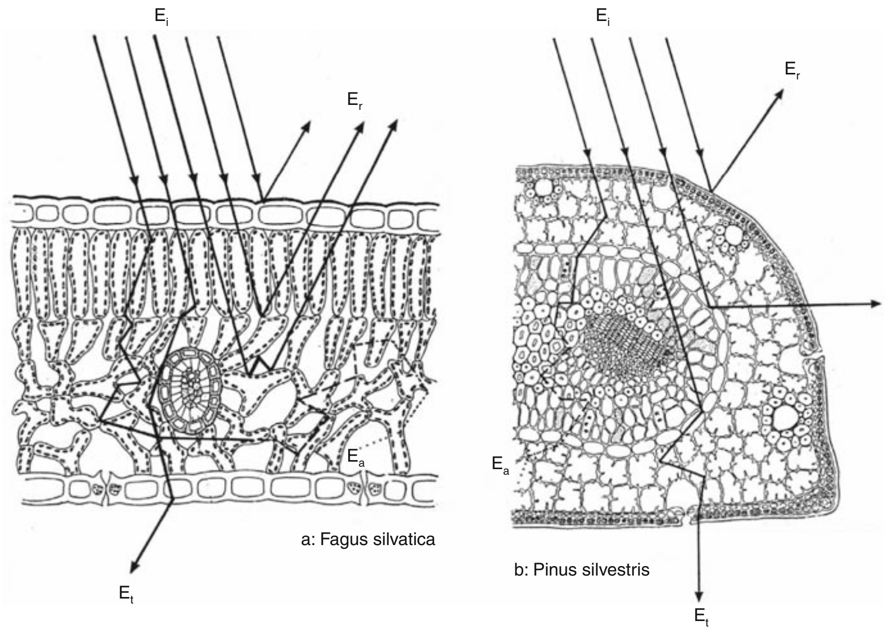


Fig. 8.56 The different spectral behaviour of broad-leaved (**a**) and coniferous (**b**) is related to the leaf structure

For the needle leaved the reflectance is lower, around 30–40% in pine trees (*Pinus spp.*) and 20–25% in fir trees (*Abies spp.*).

The spectral behaviour of a similar crop type, i.e. rice, *Oryza sativa* L., varies during the phenological cycle and is influenced by many factors that can affect the reflectivity like sowing density, variety, cultural and agronomic practices (manuring and flooding), soil type, latitude, etc.

8.3.2.6 Soil Spectral Behaviour

Remote sensing applications for soil study, survey and classification are based on some essential elements that determine their spectral behaviour.

The soil has specific characteristics related to its nature and origin and to environmental factors influencing its composition. Soil spectral characteristics depend on *chemical parameters* (organic matter, iron oxides, solvable salts, carbonates) and *physical parameters* (moisture, texture, structure, morphology).

Chemical parameters influence the soil's hue, which is darker in the presence of humus and iron oxides and lighter if calcium carbonates, silica and aluminium compounds are dominant.

In particular the humified organic matter content is decisive for the colour:

- the *humus* with high humic acid content has low reflectance values (2–5%), and it is responsible for the grey-black colour of soils with high absorption in near and medium infrared bands;
- *fulvic acids* characterize soils with the highest reflection coefficients, however not uniform at the different wavelengths. In particular the organic matter affects the soil reflectance in the 0.70–0.75 μm spectral range.

The physical parameters texture, structure and morphology do not affect the soil spectral behaviour, but the colour intensity, from the light tones to the darker ones according to the increasing moisture content.

Soil origin influences its mineralogical content, which is related to the parent material; physical–chemical processes that the soil undergoes affect its mechanical characteristics: granulometry, permeability, colour, etc. The soil elementary particles form with the organic matter different shape and size organic aggregates, contributing to the reflection behaviour.

A humid soil has a low spectral reflectance coefficient and appears darker than a dry soil; decrease/increase of soil moisture does not alter the typical spectral response of that soil, but has an influence on a higher or lower reflectance at the different wavelengths (Fig. 8.57). In a soil the moisture affects the spectral reflectance curves characteristics in particular in the near infrared: higher moisture determines a progressive irradiance decrease in correspondence of the water absorption wavelengths (2.4, 1.9 and 2.7 μm), inducing a flexion of the curves proportional to the moisture content.

Spectral intervals in the range 8–14 μm are appropriated for lithological studies, in the thermal infrared region characterized by minimal atmospheric absorption and

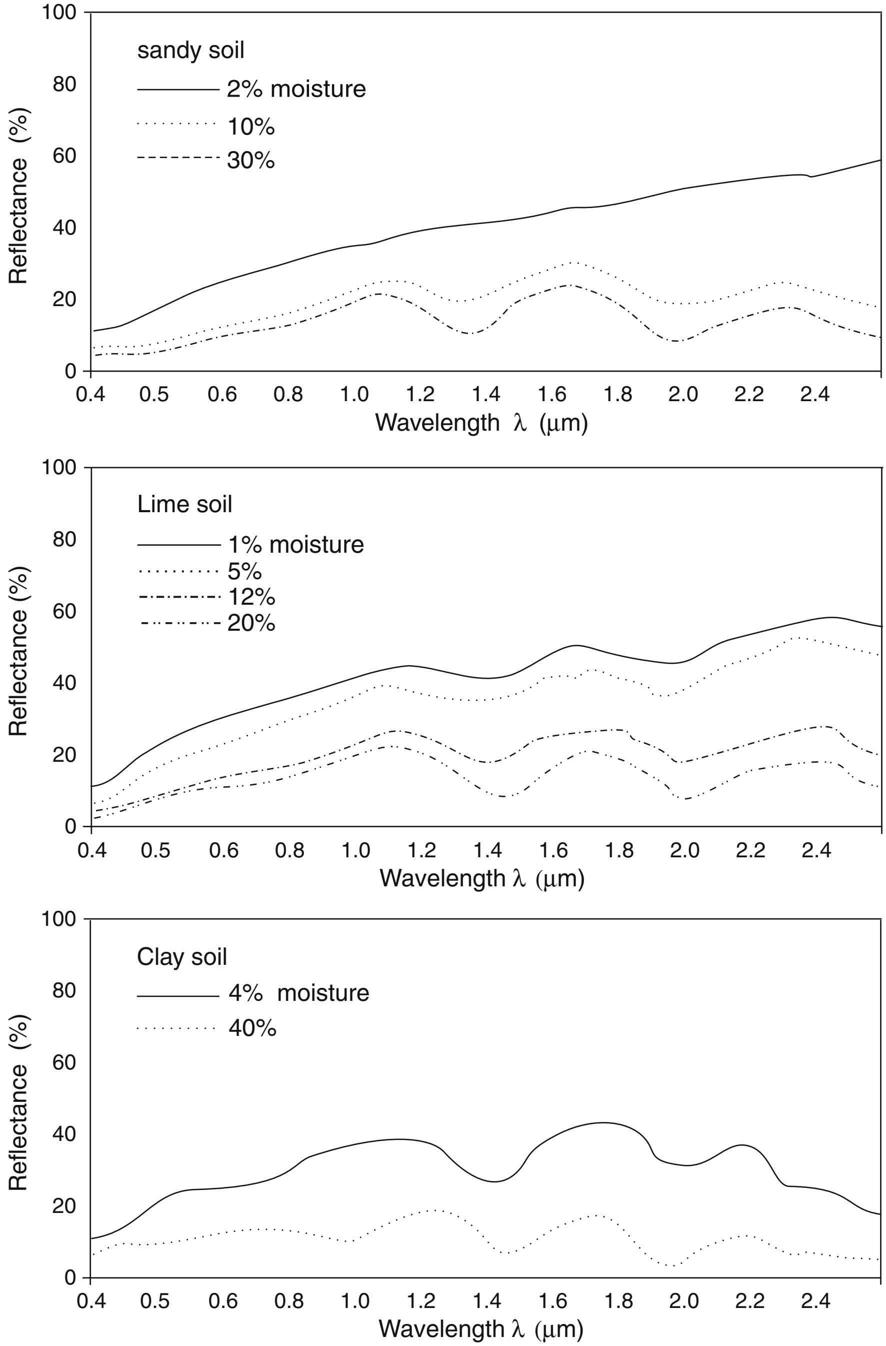


Fig. 8.57 Spectral responses of different soil types with different moisture

maximum propagation of the radiant energy. The discriminating element at these wavelengths is the thermal inertia. The temperature of a material with low thermal inertia changes significantly during the day, while the temperature of a material with high thermal inertia does not change as drastically.

8.3.2.7 Water Spectral Behaviour

Natural waters, both fresh and saline, are a witch's brew of dissolved and particulate matter whose variations in kind and concentration can indicate either pollution processes or good quality status. When deterioration of water quality is caused by optically active substances, the effect of these changes can be observed with optical remote sensing instruments.

With respect to land surfaces, water bodies are spectrally distinguishable beyond the near infrared wavelengths (>0.8 μm) since water absorbs all the incident energy, while land considerably reflects part of it. For assessing concentrations of the optically active substances, the visible part of the electromagnetic spectrum (0.4–0.7 μm) is instead the most relevant.

Remote sensing of water quality is based on simple concepts. Sunlight, whose spectral properties are known, enters a natural water body. The spectral character of the sunlight is then altered, depending on the absorption and scattering properties of the water body, which of course depend on the types and concentrations of the various constituents of the particular water body. For clear waters the water leaving radiance is higher in the blue (0.44 μm), medium in the green (0.55 μm) and negligible in the near infrared (– 0.75 μm). However, this is a weak signal; usually lower than 10% (and often less than 1%) of the total radiance measured by the sensor. Therefore, when a signal coming from the water column must be detected from space, the observing sensor requires a high signal-to-noise ratio.

In natural waters, there is a homogeneous combination of four basic components influencing their optical properties:

- pure water;
- phytoplankton;
- non-algal particles;
- coloured dissolved organic matter or yellow substances.

Pure water: regarding natural water bodies, the water is chemically pure; it is a mix of different hydrogen and oxygen isotope molecules. The smoothing beyond 0.58 μm wavelength is due to molecular absorption. The water absorbs the incident radiation in the red interval, and the absorption becomes even higher in the near infrared range at wavelength higher than 0.75 μm. In the wavelength interval between 0.4 and 0.52 μm, instead, the scattering process is dominant. The blue colour of seawater without suspended material derives from the dominant scattering of the radiation with wavelength values in the visible near to 0.4 μm and from the molecular absorption in longer λ. Some absorption, scattering and smoothing coefficient values of pure water are listed in Table 8.1.

Table 8.1 Absorption a(λ), scattering b(λ) and attenuation c(λ) coefficients of the water at 20°C. These coefficients can influence changing up to 15% of the signal in the temperature range from 0.1 to 26°C

Wavelength (μm)	a (λ)	b (λ)	c (λ)
0.4	0.006	0.0048	0.0108
0.45	0.004	0.0032	0.0072
0.5	0.006	0.0019	0.0079
0.55	0.035	0.0013	0.0363
0.6	0.200	0.0009	0.2009
0.65	0.270	0.0007	0.2707
0.7	0.600	0.0005	0.6005
0.75	2.620	0.0004	2.6204
0.8	2.020	0.0003	2.0203

Phytoplankton: these ubiquitous microscopic plants occur with incredible diversity of species, size, shape and concentration. The presence of photosynthetic pigments and associated products, coming from planktonic cells and macrophytes plants' decay, can be detected from remote sensing. It has long been recognized that phytoplankton are the particles primarily responsible for determining the optical properties of most oceanic waters, while in lakes and coastal zones phytoplankton concentration can be considered a proxy of the trophic state. Pigments like chlorophylls, carotenoids (β-carotene) and phycobilines may have significant impact on the water leaving radiances. The chlorophylls and carotenoids, which are present in all the algae species, the first one prevailing, are hence interesting parameters to be detected to describe the water status. The phytoplankton specific absorption coefficient varies with phytoplankton pigments size, shape, type and concentration. Two peaks characterize the absorption due to chlorophyll: the higher in the blue (0.44 μm) and second in red (0.68 μm)

Non-algal particles: both of organic and inorganic origin, these particles may consist of finely ground quartz sand, clay minerals or metal oxides (inorganic non-algal particles) as well as by particles which are produced, for example, when phytoplankton die and their cells break apart, and when zooplankton graze on phytoplankton and leave cell fragments and fecal pellets (organic nonliving particles). The optically properties of these particles can vary considerably depending by their nature; for instance there is a relation between the increase of concentrations of inorganic particles and the increase of the radiance reflected at the longer wavelengths.

Coloured Dissolved Organic Matter: these compounds represent the coloured fraction of compounds' decayed products from plant matter and consist mostly of various humic and fulvic acids. They are also known as yellow substance, gelbstoff or gilvin. CDOM absorbs very little in the red, but its absorption increases rapidly with decreasing wavelength and can be significant at blue and ultraviolet wavelengths. One of its main sources is decayed terrestrial vegetation. Concentrations are thus generally greatest in lakes, rivers and coastal waters influenced by river runoff.

By knowing the spectral properties of water components briefly described above, the deconvolution of remotely sensed signal (inclusive also of the atmospheric contribution) allows the detection of green algae pigments, total suspended matter, coloured dissolved organic matter, of water transparency and of the Secchi disk, the last two strictly related to the diffused attenuation. When the water transparency allows the light to reach the bottom, the backward signal reaching the sensor also include spectral information of the substrate. In such cases, the spectral response of bottom and the water depth becomes detectable from remote sensors too.

The main components determining the radiance values of a natural water body recorded by a sensor are (Fig. 8.58):

L_a : radiance portion resulting from the atmospheric scattering, which does not reach the air/water surface

L_s : radiance portion reaching the air/water interface but not penetrating it

L_v : portion penetrating the surface and reflected before it reaches the water body bottom

L_b : portion penetrating the surface and reflected by the water body bottom.

Remote sensing water quality is also related to water colour and to its definition based on the CIE colour coordinate system. The general assumption is that the water colour qualitatively may describe the amount of concentration of suspended solids (generally in terms of inorganic nonliving particles), of phytoplankton and more generally the water transparency. For instance, an increase of radiance at the red wavelengths indicates decreased transparency in the water and increased suspended material while an increase of radiance at the green wavelengths is correlated to an increase of chlorophyll.

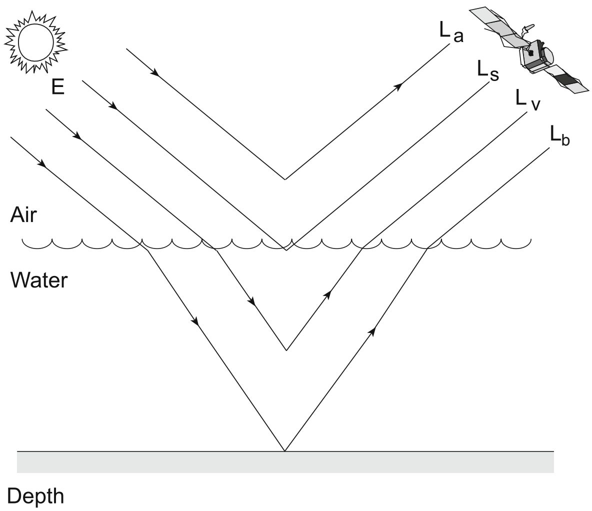


Fig. 8.58 Components of the energy fluxes that determine signal of a water body sensed by a remote sensor. The example is limited to the solar and atmospheric reflected radiation, representing the descending radiation interfering with water and atmosphere

It is interesting to compare the water spectral properties to the reflectance of snow. Contrary to the water bodies, water in its frozen states exhibits much higher reflectivity, which can be close to 1 in the visible. Snow is a collection of ice grains and air, and, when at 0°C, it also has a significant fraction of liquid water. Snow also often includes particulate and chemical impurities, as dust, soot, pollen and other plant material, and small amounts of the major anions and cations. Thus the optical properties of snow depend on the bulk optical properties and the geometry of the ice grains, the liquid water inclusions, and the solid and soluble impurities (Figs. 8.59 and 8.60).

Snow reflectance decreases at longer wavelengths becoming spectrally distinct by clouds.

8.3.2.8 Multi-temporal Images

The spectral behaviour of objects changes according with the moment of the observation (Fig. 8.51). This aspect is determinant in studying the canopies having a specific phenological cycle, such as wheat with respect to corn and soybeans, and these with respect to sugar beet. The wheat/barley and corn/soybean cycles are difficult to be distinguished because they differ only for few days, although the different morphology can be a detectable element.

Hence the use of multispectral and multi-temporal images is important.

Satellite sensors collect terrestrial surface radiance data at regular time intervals (multi-temporal) revisiting the same scene. The interval of acquisition over a same area (temporal resolution) can vary between a few hours, 12 for NOAA, to days, 16 for Landsat, 26 for SPOT and 35 for Envisat.

The cloud cover is an obstacle to the optical detection, reducing the possibilities of image choice and use. The longer the temporal resolution the lower the probability to obtain images of the same scene without clouds. The possibility of programming the acquisitions contribute to partially solve this problem.

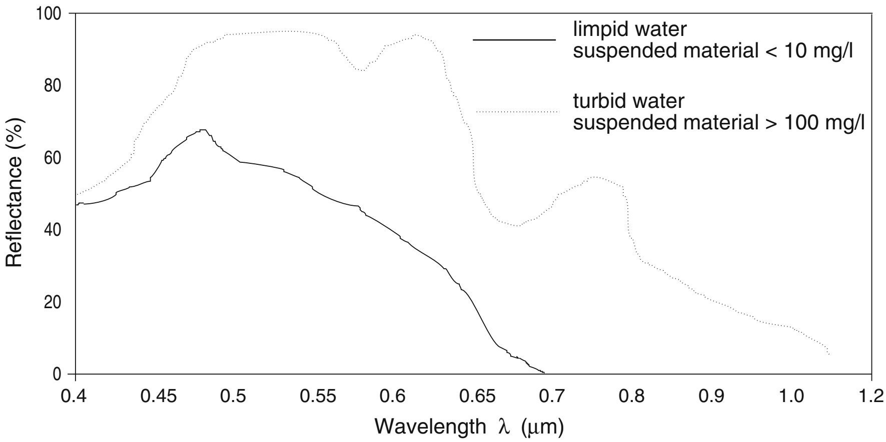


Fig. 8.59 Spectral responses of limpid water and with suspended material in the spectral range of the visible/near infrared

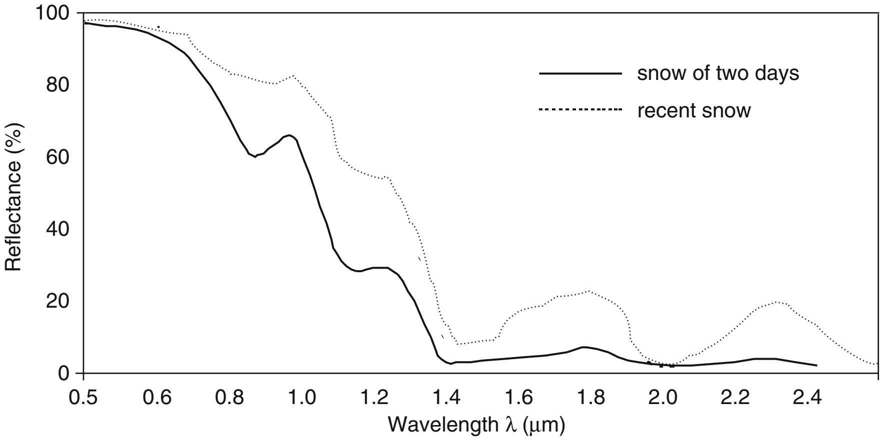


Fig. 8.60 Spectral behaviour of two types of snow

The active systems, i.e. the instruments operating in radar band frequency, have been planned to overcome this limit as the emitted signals, in the microwaves spectral domain, are little or not at all influenced by the atmospheric conditions.

The integration between passive and active systems can contribute to improve the temporal acquisitions of the same areas providing alternatively spectral and/or backscattering data according to the optical or radar sensors characteristics.

Some European projects, MARS (*Monitoring Agriculture with Remote Sensing*) and *CORINE Land Cover* (CLC), have largely used remotely sensed data for the land use/land cover classification. MARS project adopted semi-automatic classification systems with the use of parametric and non-parametric classifiers; CORINE Land Cover is based on the visual interpretation of multi-temporal images (EC, 1993). FAO Africover project uses a dichotomous method, based on the use of an expert system enabling a better classification flexibility and dynamicity starting from satellite image multispectral and multi-temporal analysis (see Chapter 10).

8.4 Quantitative Analysis

8.4.1 Multispectral Transformation of the Images (Vegetation Indices)

The necessity to observe the natural and human territory conditions over a time period that can vary from months up to many years requires the use of indices uniform and directly comparable among them.

Vegetation Indices (VIs) are related to vegetation cover and status and provide useful information on biomass productivity and health condition.

VIs have a direct correlation with leaf chlorophyll content and *Leaf Area Index* (LAI) and vary in relation to vegetation cycle and phenology.

They are also sensitive to other external factors:

- contribution of the soil/background optical behaviour where the vegetation does not completely cover the ground;
- geometry of view due to sensor angles' acquisition and to Sun position, as well as the atmospheric effects, treated in the Chapter 4 about geometric and radiometric corrections;

There are two different approaches to define Vegetation Index:

- through the conservation and combination of the original radiometric data, namely VIS and NIR bands diagnostic of vegetation spectral response;
- through modification of the available original spectral bands acquired by the same sensor, like the *Kauth–Thomas Tasseled Cap transformation*.

Vegetation Indices not modifying the original radiometric information enhance the remarked difference between the low reflectance in the visible, in particular in the red and blue, and the high answer in the near infrared. The combination can be in the form of ratio, slope or other formulations (see Chapter 4).

These indices can be divided in four categories (Box 8.9):

Box 8.9 Classification of the vegetation indices (VIs)

No radiometric modification	Intrinsic	Operation between bands
		NDVI
		SAVI NBR BAI UI SVI
	In relation with the soil line	PVI TSAVI MSAVI
	With atmospheric corrections	ARVI GEMI
Radiometric modification (pseudo-bands)	Same number of bands	Tasseled Cap Principal Component Analysis (PCA)
	Reduction of number of bands	Principal Component Analysis (PCA) Canonical Analysis

- *intrinsic indices*, like the *Simple Ratio* (SR), also called *Ratio Vegetation Index* (RVI), arithmetic operations between bands (Pearson & Miller, 1972) and the *Normalized Difference Vegetation Index* (NDVI), based only on the spectral reflectance measured;
- *indices related to the soil line*, introducing the use of parameters related to the soil, like the *Perpendicular Vegetation Index* (PVI) and its derived *Weighted Difference Vegetation Index* (WDVI);
- soil-adjusted Indices as *Soil-Adjusted Vegetation Index* (SAVI) and its modifications TSAVI, MSAVI;
- *indices with atmospheric corrections*, like the *Atmospherically Resistant Vegetation Index* (ARVI) and the more complex *Global Environment Monitoring MSAVI* (GEMI) (Box 8.12).

Furthermore, VIs can be divided according to the slope (NDVI) or on the distance (PVI) from the soil line (Fig. 8.61).

The development of hyperspectral sensors allows deriving a set of indices calculated exploiting the information of narrow bands in diagnostic vegetation spectral feature such as *Red Edge Position* (REP), the transition zone between R absorption

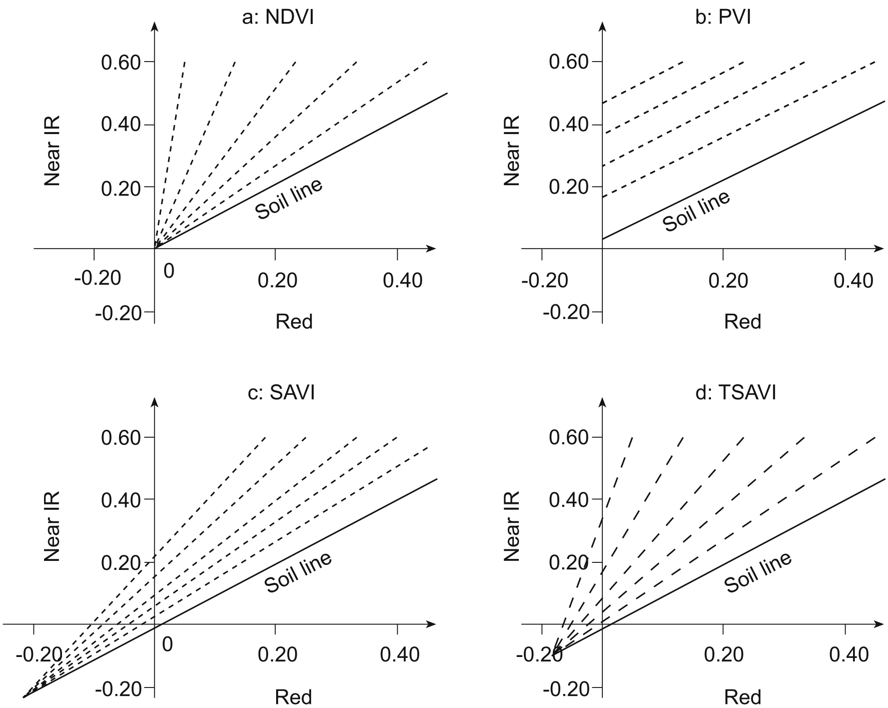


Fig. 8.61 Vegetation isolines concept for different vegetation indices: (**a**), (**c**) ND (**d**) are linked to the slope and (**b**) based on the distance (Qi et al., 1994)

and NIR reflection. This category of indices can be derived directly from bands calculation or studying the spectral response. Hyperspectral sensors provide a quasi-continuum description of vegetation spectra that can be characterized calculating minimum, maximum and flex position of the spectral responses.

8.4.1.1 Indices with Conservation of the Original Radiometric Data

Operation Between Bands

The easiest computation of image indices conserving the original data is based on subtraction, division, sum and multiplication operations of homologous cells' radiance values (expressed in watt per surface unit) or pixels' brightness (expressed in Digital Numbers) of two different spectral bands acquired by the same instrument to produce a new image; the most used is subtraction and division.

These algebraic combinations involving two or more spectral bands provide information on plant status and health; multi-temporal acquisition of Earth Observation data permits vegetation continuous monitoring.

In particular, according to the ratio between bands, each pixel's DNs of an image (spectral band 1) is divided by the DN of the correspondent pixel of another image (spectral band 2) that has been recorded on the first one (Box 8.10).

Box 8.10 Band ratios and element emphasized. NIR: near infrared, MIR: medium infrared, R: red, G: green

Ratio	Enhanced elements
$\frac{NIR}{R}$ $\frac{NIR}{V}$	Vegetation, foliar biomass
$\frac{V}{R}$ $\frac{V}{IR}$ $\frac{R}{IR}$	Soil, rocks
$\frac{MIR}{V}$ $\frac{MIR}{R}$	Soil moisture, water stress of vegetation

Multispectral image ratio enhances spectral reflectivity small differences between the surfaces of materials which would be hard to distinguish on the original data (Fig. 8.62). Moreover this removes first-order albedo variations due to the topography: i.e. a spectral data normalization is obtained eliminating the illumination contrasts and enhancing the radiometric data content.

The first information about the vegetation spatial distribution and its monitoring comes from the *Simple Ratio* between the reflectance in the near infrared (NIR: 0.76–0.90 μm) and in the red (R: 0.63–0.69 μm) (Figs. 8.63 and 8.64a).

Normalized Difference Vegetation Index (NDVI)

A particular role is covered by the *Normalized Difference Vegetation Index* (NDVI), which has the advantage to allow comparisons between images acquired in different times (Rouse et al., 1974). Also the NDVI, like all the vegetation indices, relates the

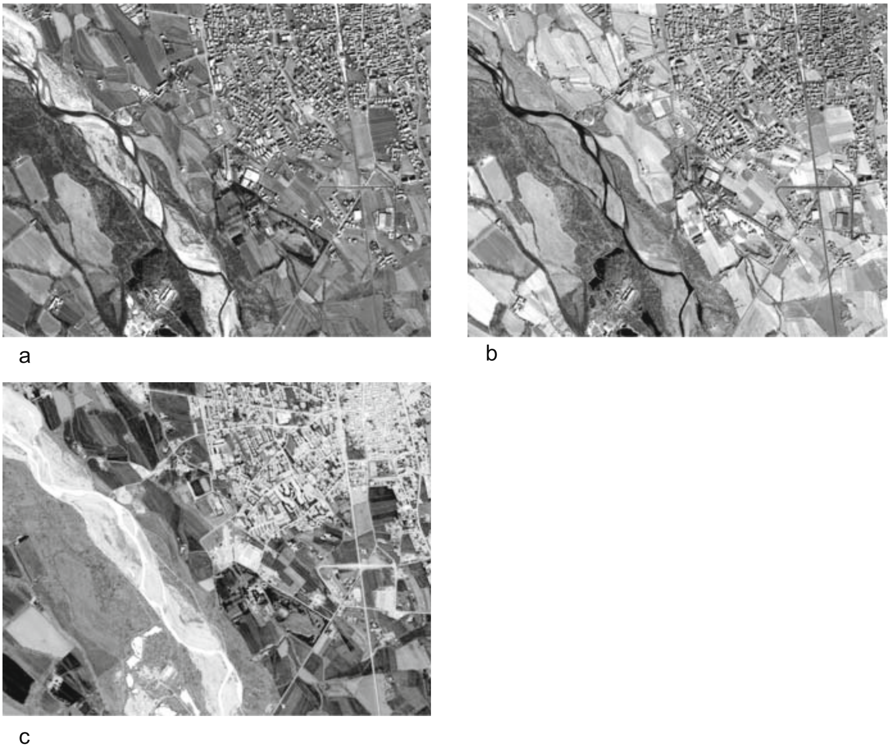


Fig. 8.62 Ratio between the TM3 red (a) and TM4 infrared (b) spectral bands. The difference between vegetation (light tones), bare soil and water (darker tones) is emphasized in the ratio

spectral absorption of the chlorophyll in the red with a reflection phenomenon in the near infrared, influenced by the leaf structure type.

The NDVI value is defined as:

$$NDVI = \frac{\rho_{IR} - \rho_R}{\rho_{IR} + \rho_R} \tag{8.10}$$

where

ρ_{IR} = reflectivity in the near infrared band
ρ_R = reflectivity in the red band.

This function, ratio between the difference and the sum of two bands, exists in the interval between −1 and +1. The actual values included between −0.1 and +0.6 are related to the vegetation physiological conditions and to the biomass level.

Figure 8.64b shows the equivalence between the Vegetation Index NDVI and the ratio NIR/R with a 45° rotation with respect to what is represented in Fig. 8.64a. The *soil line*, in Fig. 8.64b, is represented by the diagonal with origin in the axes and

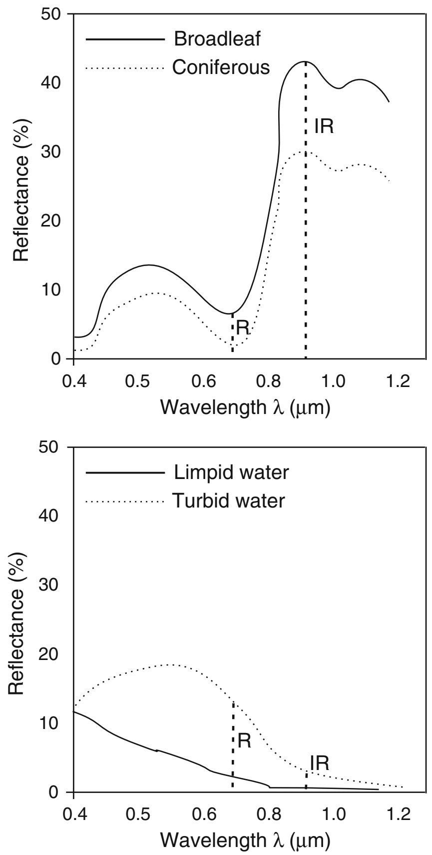


Fig. 8.63 Radiance values in the red R and near infrared (IR) for vegetation, soil and water. The ratio near infrared/red (IR/R) is positive and very high for the broad-leaved and positive and high for the coniferous, slightly positive for the soil, negative for the water

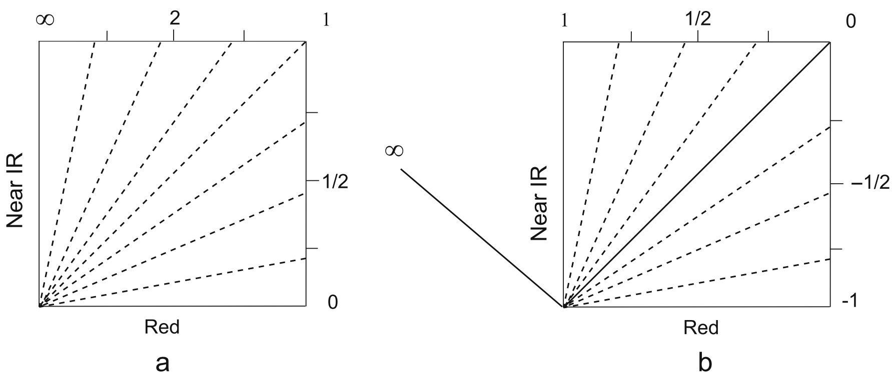


Fig. 8.64 (a) Isolines of the ratio IR/R. The index values are in the range from 0 to infinite (∞); (**b**) isolines of the Normalized Difference Vegetation Index (NDVI) with values ranging from -1 to $+1$

culminates in the value 0. The abscissa is identified as the ground axis, uniformly equal to 0, while vegetation values, reported on the y-axis, vary between 0 and 1.

The simplification of making the red band coincide with the soil line derives from the fact that at 0.63–0.69 μm wavelength, red one's, vegetation reflection is reduced and the values are relatively higher for the soils. To coincide with the x-axis means to have reduced (positive values near to the 0: soil) or null (negative values: water) answer in the near infrared, hence vegetation absence. This concept, represented also as soil line, i.e. complete vegetation absence, is used also in other algorithms defining vegetation indices.

The axes' rotation generates negative values which can be compensated by the computation (8.11). As the values referred to the soil are distributed along the diagonal, the negative values correspond to the class water from clear to different turbidity levels.

To simplify the image processing, (8.10) has to be transformed into an entire function whose real numbers are transformed into Digital Numbers in an interval ranging between 27 and 227, 8 bit scale with 256 (0–255) levels, with a 0.01 equivalent increase for each DN:

$$NDVI = \frac{\rho_{IR} - \rho_R}{\rho_{IR} + \rho_R} \cdot 100 + 127 \tag{8.11}$$

Other Indices

Intrinsic indices generate values difficult to interpret when the biomass is low and the soil nature is unknown. The indices related to the soil line are based on the principle of linearity of the soil's spectral behaviour in the visible and near infrared, consequently deriving some transformation coefficients (Box 8.11).

For example the PVI expresses the distance between the biomass reflectance in the red and near infrared and the soil line. Also the PVI is negatively affected by the presence of a low vegetation density, while the SAVI and the following modified versions (TSAVI, MSAVI) bring significant improvements.

Huete (1988) developed the SAVI by shifting the convergent point of iso-vegetation lines from the origin to a point in the quadrant of negative R and NIR values (Fig. 8.61). Many studies have indicated that SAVI not only reduces the effect of soil variability for low LAI, but also increases the sensitivity to high LAI (Elvidge & Chen, 1995).

8.4.1.2 Indices with Modification of the Radiometric Information

Remotely sensed images' multispectral character makes possible to generate new bands, called pseudo-bands, derived from the original ones also called components. Objects otherwise not visible in the original scenes can be so far discriminated.

The components represent an alternative description of the original bands generally obtained as their linear combination. Such kind of multispectral transformations are the Tasseled Cap Transformation.

Box 8.11 Definition, acronym and algorithm of some vegetation indices

Vegetation index	Acronym	Algorithm
Perpendicular Vegetation Index (Richardson & Wiegand, 1977)	PVI	$\frac{(\rho_{NIR} - a\rho_R - b)}{\sqrt{1+a^2}}$
Soil- Adjusted Vegetation Index (Huete, 1988)	SAVI	$\frac{(\rho_{NIR} - \rho_R)}{\rho_{NIR} + \rho_R + \mathrm{L}}(1+\mathrm{L}) \quad \mathrm{L} = 0.5$
Transformed Soil-Adjusted Vegetation Index (Baret & Guyot, 1991)	TSAVI	$\frac{a(\rho_{NIR} - a\rho_R - b)}{\left[\rho_R + a(\rho_{NIR} - b) + 0.08(1+a^2)\right]}$
Modified Soil-Adjusted Vegetation Index (Qi et al., 1994)	MSAVI	$\frac{(\rho_{NIR} - \rho_R)}{\rho_{NIR} + \rho_R + \mathrm{L}} \quad (1+\mathrm{L}) \quad \mathrm{L} = 1 - 2a \cdot NDVI \cdot (\rho_{NIR} - a\rho_R)$ a: slope of the soil line
Atmospherically Resistant Vegetation Index (Kaufman & Tanré, 1992)	ARVI	$\frac{(\rho_{NIR} - \rho_R \cdot \rho_B)}{(\rho_{NIR} + \rho_R \cdot \rho_B)} \quad \rho_R \cdot \rho_B = \rho_R - \gamma(\rho_B - \rho_R)$ γ: type of aerosol
Global Environment Monitoring Index (Pinty & Verstrate, 1992)	GEMI	$\eta(1 - 0.25\eta) - \frac{(\rho_R - 0.125)}{(1 - \rho_R)}$ $\eta = \frac{2(\rho_{NIR}^2 - \rho_R^2) + 1.5\rho_{NIR} + 0.5\rho_R}{(\rho_{NIR} + \rho_R + 0.5)}$
Normalized Burn Ratio (Key & Benson, 1999)	NBR	$\frac{\rho_{IR} - \rho_{SWIR}}{\rho_{IR} + \rho_{SWIR}}$
Burned Area Index (Martin 1998)	BAI	$\frac{1}{(\alpha - \rho_R)^2 + (\beta - \rho_{IR})^2}$ with $\alpha = 0.1$ and $\beta = 0.06$
Urban Index (Kawamura et al., 1996)	UI	$\frac{\rho_{SWIR} - \rho_{NIR}}{\rho_{SWIR} + \rho_{NIR}}$
Soil and Vegetation Index	SVI	$\frac{\rho_{SWIR} - \rho_B}{\rho_{SWIR} + \rho_B}$

Tasseled Cap Transformation (TCT)

Kauth & Thomas transformation (1976), also called Tasseled Cap, successfully applied to Landsat TM images (Crist & Cicone, 1984), proposes to generate synthetic bands with decorrelated information content and specific physical meaning starting from the original information collected by the satellites (Fig. 8.65):

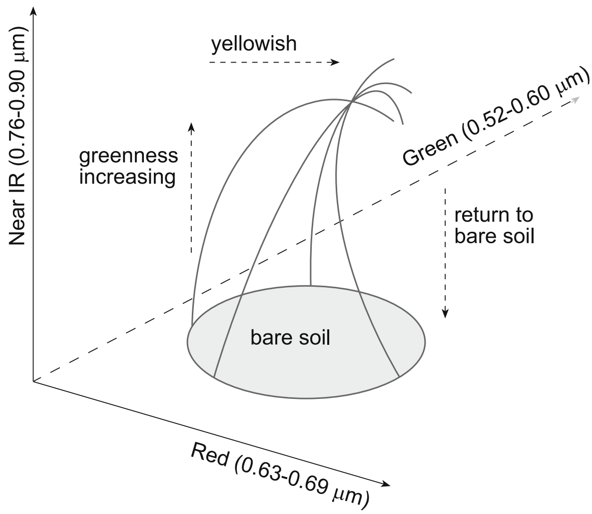


Fig. 8.65 The Tasseled Cap transformation summarizes the temporal modification of the spectral reflectance of a crop cycle. Three orthogonal directions are identified: the principal diagonal, along which soil are distributed, is represented by the axis of the soil brightness; the development of the biomass to a maximum follows the greenness axis as crops move towards maturity (green stuff); crop yellowing takes place in an orthogonal axis to the soil line and return to the bare soil response (yellow staff) (Kauth and Thomas, 1976)

- *brightness* index, indicating the radiance intensity, related to reflectance and albedo values (reflectance energy from the surfaces in the visible and IR) of non-vegetated area;
- *greenness* index, enhancing the contrast between the visible and infrared (IR) region bands; indicator of the quantity and, at some conditions, of the quality of the biomass;
- *yellowing* index (*yellow stuff*) that, for example for the herbaceous crops, detects the ripening and harvesting phases and the bare soil condition after those;
- *wetness* index, sensitive to the contrast between the medium-near IR and visible-near IR regions, providing information about soil moisture and biomass (applicable on Landsat TM and ETM+) Table 8.2.

These units help to understand in spectral terms the soil–vegetation characteristic relation according to vegetation temporal changes. Chlorophyll content variations have a characteristic behaviour that can be highlighted by temporal and spectral graphic representation of greenness index.

Hence vegetation indices, in accordance with the most relevant experiences carried out in the last years, can be actually used in the following applications:

- plant biomass study and assessment in general;
- definition of forest wood degradation due to logging and fires, also in relation to the effects of soil conservation;

Table 8.2 Synthetic bands derived from the Tasseled Cap transformation of a Landsat MSS image, bands 4, 5, 6, 7: blue-green, red, near infrared, near infrared. Variable correction factors are defined for each band and index in function of environmental parameters

Synthetic band	Algorithm
Soil Brightness Index (SBI)	0,3328 · b4 + 0,603 · b5 + 0,675 · b6 + 0,262 · b7
Green Vegetative Index (GVI)	− 0,283 · b4 − 0,660 · b5 + 0,577 · b6 + 0,388 · b7
Yellow Staff Index (YVI)	− 0,899 · b4 + 0,428 · b5 + 0,076 · b6 · 0,041 · b7
Non-Such Index (NSI)	− 0,016 · b4 + 0,131 · b5 - 0,452 · b6 + 0,882 · b7

- monitoring damages to crops due to infestations, pathologies and water stress;
- evaluating and monitoring phenological phases for agro-meteorological models ideation.

Feature Reduction

Feature reduction intends to synthesize the information contained in the original data and can be used to concentrate the information in a smaller number of bands. This can be useful to reduce the amount of data to be processed, to transmit or to store concentrating, for example, the most part of the original data information on a small number of synthetic bands. Features which do not aid discrimination and separability of spectral classes can be eliminated. The Principal Component Analysis (PCA) and the Canonical Analysis are the typical algorithms used.

In the case of a two-dimension PCA transformation (to simplify), the new generated bands result from the x, y axes rotation with fulcrum in their origin (Fig. 8.66). An image cell is defined as a point in the bands' hyper-space; it can be represented as a vector where each component is represented by the DN of one of the bands. The new pixels radiance values, linearly related to the original ones, are obtained

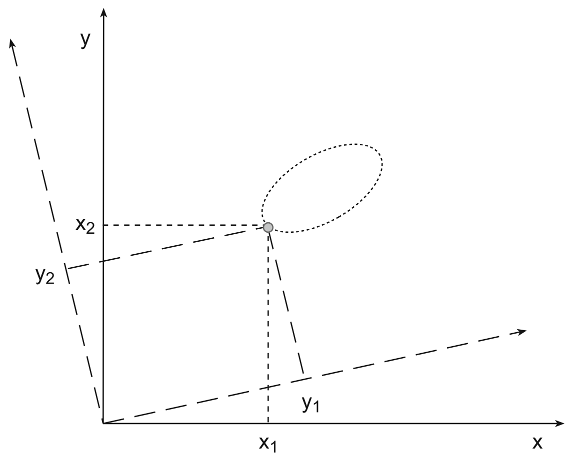


Fig. 8.66 Modified coordinates system where the pixel's new vectors can produce a data decorrelation, increasing the separability among pixels classes

by using geometrical–statistical units, such as the auto-vectors, the mean and the covariance matrix.

8.4.2 Classification Techniques

To better use the information contained in spectral bands, procedures that recognize the spectral homogeneity, based on properly developed algorithms, are applied (Box 8.12).

Through the classification, classes are recognized on the image based on spectral characteristics (Fig. 8.67). The classifiers are software instruments that perform this operation (Box 8.13) and can be

Box 8.12 Fundamental definitions and indices for land use/land cover analysis. The indices are assessed using various approaches and algorithms applied to many sensors data

Parameter	Definition
Land-Surface Reflectance	Part of solar radiation reflected by the land surface. The measured surface reflectance depends on the Sun zenith angle and on the viewing angular configuration. Two successive measurements of the surface reflectance cannot be directly compared The directional effects have to be removed using a normalization algorithm before generating a composite
Land Cover map	Associates to each pixel of the surface a labelling characterizing the surface (ex : deciduous forest, agriculture area, etc.) following a pre-defined nomenclature
Bi-directional Reflectance Distribution Function (BRDF)	Gives the reflectance of a target as a function of illumination geometry and viewing geometry BRDF depends on wavelength and is determined by the structural and optical properties of the surface
Leaf Area Index (LAI)	Half the total foliage area per unit of ground surface (Chen and Black, 1992)
fAPAR	Fraction of photosynthetically active radiation absorbed by vegetation for photosynthesis activity. The fAPAR can be instantaneous or daily
Fcover	Fraction of ground surface covered by vegetation

- *parametric*, based on distribution models defined by geometric–statistical parameters (distance, mean, standard deviation, etc.);
- *non-parametric*, based on operative models whose internal parameters do not have a geometric–statistical value; they are used in cases of uncertainty representation of the spectral distribution like in the case of multi-sensor data classification.

The choice between parametric and/or non-parametric classifier is based on the following criteria :

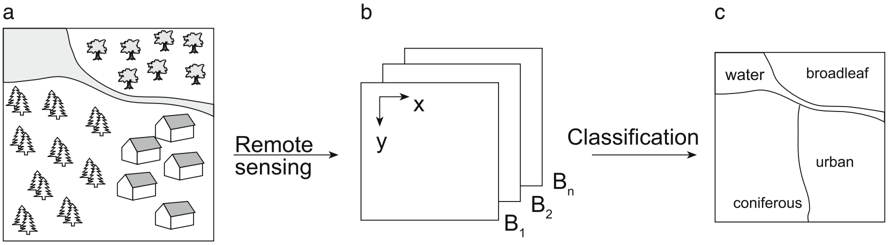


Fig. 8.67 From the real world (**a**) more spectral images can be obtained (**b**) to be used for a classification process and thematic maps production (**c**) based on selected legends

Box 8.13 Classification of the image classifiers

Classification		
Unsupervised	Supervised	
	Parametric classifier	Non-Parametric classifier
Mobile means (K-means)	Maximum likelihood	Fuzzy set
Isodata	Minimum distance	Neural network
	Parallelepiped	Change detection

- computation costs (training, classification);
- (potential) accuracy, assessed according to the data (training data) available for the classification and to the classification model capacity to adapt to the available data;
- other costs, such as hardware, ground truth data collection and operators/analysts time.

Classifiers can be distinguished into:

- *unsupervised*, or automatic, classification method does not depend on the knowledge of the ground truth and does not require the availability of external information for assigning the pixels to the different classes;
- *supervised*, or semi-automatic, classification is more accurate than unsupervised for mapping classes, but depends on the cognition and skills of the image specialist. The strategy is based on a priori knowledge of the classes for a sufficient number of pixels (training sets) that the analyst must recognize by personal experience, experience with thematic maps, and in-field survey. The left pixels, in a much higher number, are then classified according to their similarity with training sets.

Many assumptions are typically made in the course of a supervised digital image classification. This assumption of an exhaustively defined set of classes is not always

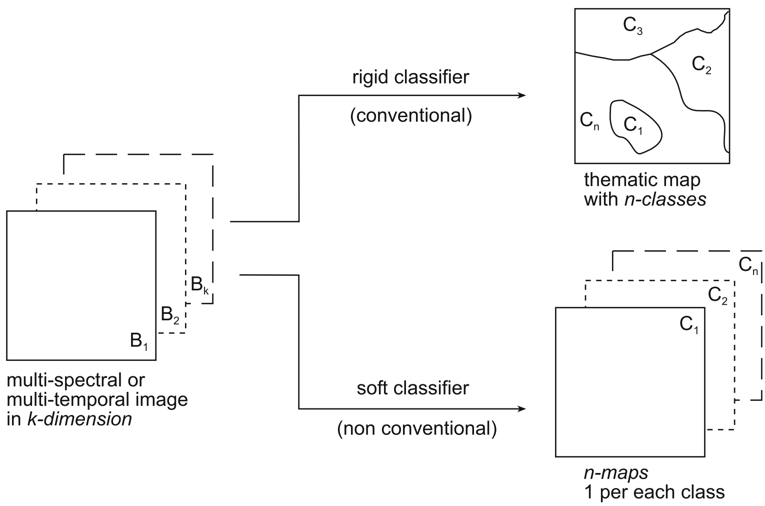


Fig. 8.68 The difference between a conventional hard classifier and a soft classifier

satisfied, with the imagery containing regions of classes that were not included in training the classification. Different approaches can be used in classification:

- *hard classification* univocally assigns each pixel of the scene to one only class. There is no space for considerations about mixed pixels (radiometry coming out from the contribution of many land cover classes) and the result is a single image corresponding to the classification;
- *soft classification* can provide a better and more accurate representation of both discrete and continuous land cover classes, resolving in particular problems associated with mixed pixels (Fig. 8.68).

Soft classification is based on the concept of pixels belonging (or assignation probability) to the different classes. For each class an image is generated, which defines, for each cell, its belonging degree to the class to which it refers. Soft classification results can be converted into hard classification assuming as each pixel assignation class the one which is assigned the higher belonging degree.

In reality such a simple distinction is not always observed and a continuum of classification softness can be defined. Continuum classification methods are represented by neural network algorithms.

8.4.2.1 Unsupervised Classification

Unsupervised classification is used to cluster pixels in a data set based on statistics only, without any user-defined training classes. Although the method requires no user input to create the classified image, the output tends to require a great deal of post-classification operations to make the results more meaningful.

Moreover neither the nature nor the exact number of classes is supposed to be known a priori. In this situation, the only possible approach is to generate families (clusters) of radiometrically homogeneous pixels (cluster according to a reciprocal proximity criterion). Cluster purpose is grouping pixels in a priori unknown numbers of spectral classes. This purpose can be reached through different levels of difficulty and success, according to the pixels' position in the bands space.

Clustering and separation general criteria are generally based on the concept of distance that, if not specified in another way, is intended as Euclidean:

$$\rho(x',x) = ||\mathrm{x}' - \mathrm{x}||$$

It is isotropic and homogeneous in the whole multispectral space.

In this process, clusters defined by their spectral response are automatically recognized (Fig. 8.69). The procedure consists of two phases:

- definition of the number of centroids, around which pixels with similar spectral characteristics will tend to group. The initial centroids are chosen based on statistic criteria, without having a priori information about their correspondence with the land cover classes;
- each cluster is assigned to a class, referring to a relevant category of sites (soil, vegetation, water, etc.).

Mobile Means (K-Means)

K-means (MacQueen, 1967) is one of the simplest unsupervised learning algorithms. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids should be placed

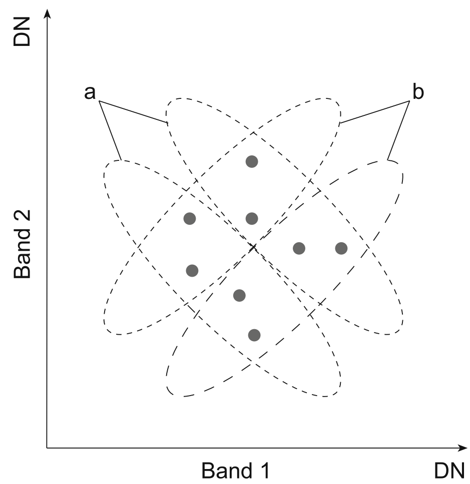


Fig. 8.69 Two possible cluster (pixels family) aggregation in a bi-dimensional space (scatterogram) that can produce very different classification results

in a cunning way because different locations cause different results. So, the better choice is to place them as far as possible from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early groupage is done. At this point, new k centroids are re-calculated as barycentre of the clusters resulting from the previous step. With these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid, generating a loop. As a result of this loop the k centroids change their location step by step until no more changes are done (Fig. 8.70).

Isodata

Isodata algorithm is derived from the *K-means*. It differs from the *K-means* as the clusters number can vary between a minimum and a maximum defined by the user. In other words, the ISODATA algorithm allows for a different number of clusters while the k-means assumes that the number of clusters is known a priori.

Isodata stands for *Iterative Self-Organizing Data Analysis Techniques*. This is a more sophisticated algorithm which allows the number of clusters to be automatically adjusted during the iteration by merging similar clusters and splitting clusters with large standard deviations. This iterative technique for choosing a threshold was developed by Ridler and Calvard (1978).

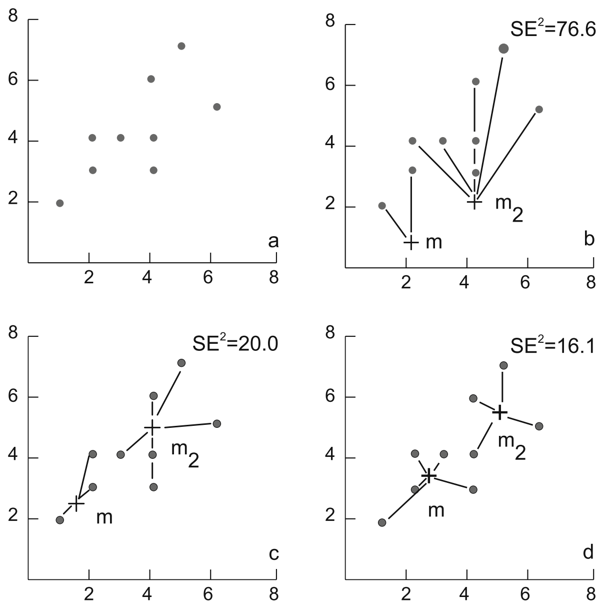


Fig. 8.70 Interactive optimization of two spectral classes (cluster m_1, m_2) for a same set of pixels (**a**) with progressive reduction of the sum of squared errors SE^2 (**b**, **c**, **d**)

The variability is ruled by the following criteria verified at the end of each interaction:

- elimination of clusters considered not significant because of the small number (to be defined) of pixels;
- grouping of the clusters with close centroids to justify the existence of one only cluster;
- splitting of over extended clusters into a higher number of new clusters.

The cluster number can decrease by eliminating or joining them as well as increase by splitting them.

The parameters defined by the user are

- initial number of clusters;
- maximum admitted number of clusters;
- minimum number of pixels assigned to a cluster;
- maximum number of interactions;
- minimum accepted distance between centroids of two adjacent clusters;
- maximum number of clusters that can be merged in each single iteration;
- maximum admitted dispersion value (measured as standard deviation) of a cluster.

8.4.2.2 Supervised Classification

In supervised classification, the imagery is divided into training data and test data. The correct categories are known for the training data, and some classification approach is specified based on this data. This approach is then used to classify the test data. Some approaches include classification trees, minimum distance statistical approaches and neural networks. The choice of a good training set can have significant influence on the success of a classification approach. A common technique in imagery classification is selection of good test data points by an experienced analyst.

The similarity concept is extended to the used spectral bands, and the proximity in the multispectral space is evaluated (Fig. 8.71). Different distance measures generate specific supervised classification algorithms, which differ in terms of computation complexity. The most effective algorithms require longer computation time but there are cases in which the easiest algorithms provide even the best results.

An essential element in the supervised classification is the definition of the classes to which the pixels should be assigned. The identification of the training sets in the image is the most critical factor for a successful classification.

Training sets are a number of pixels of a sample areas representative of cover category, known and localized on the images, used for the classification of the whole scene. In this process significant classes are selected and for each training set statistical parameters of the DNs for each band are calculated. More than one poly-

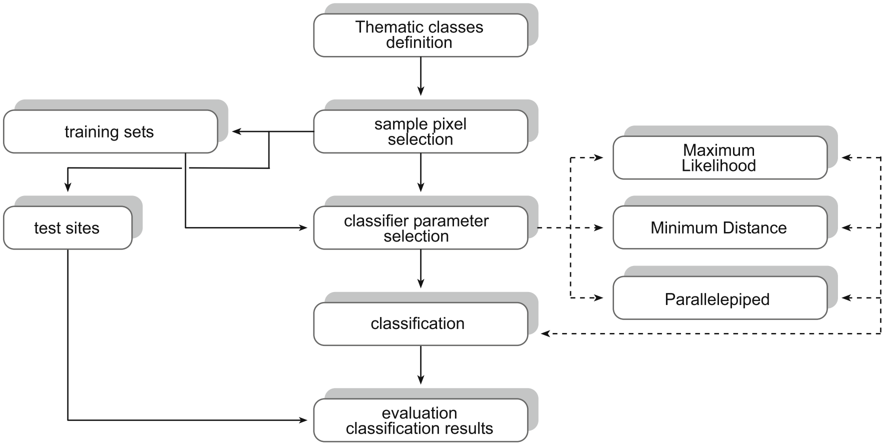


Fig. 8.71 Supervised classification phases

gon can be established for any class. When DNs are plotted as a function of the band sequence, increasing with wavelength, the result is a spectral response for that class.

The next classification step then acts to cluster the data representing each class. When the DNs for a class are plotted as a function of the band sequence, increasing with wavelength, the result is a spectral response for that class. The multiple spectral responses so obtained are for all of the objects within the site that interact with the incoming radiation. Classification then proceeds by statistical processing in which every pixel is compared with the defined various signatures and assigned to the class with the closest signature (Figs. 8.72 and 8.73). A few pixels in a scene do not match and remain unclassified, because these may belong to a class not recognized or defined.

The last step is the representation of the different classes, assigning the colour codes or conventional values, producing thematic maps.

Maximum Likelihood

In the supervised and parametric classification, it is possible to use different algorithms defining the assignation criterion of each single pixel to a class. The *maximum likelihood* (*MLL*) classifier is the most used. It operates assigning each pixel to the class for which the conditional probability, defined as the probability that, selecting a pixel x of the scene, it belongs to a certain class C, is higher. The pixel belongs to the class C if:

$$P(C|x) > P(C_k|x) \tag{8.12}$$

where $P(C \mid x)$ is the conditional probability of the pixel x with respect to the class C; $P(C_k \mid x)$ is the conditional probability of the pixel x with respect to the class C_k.

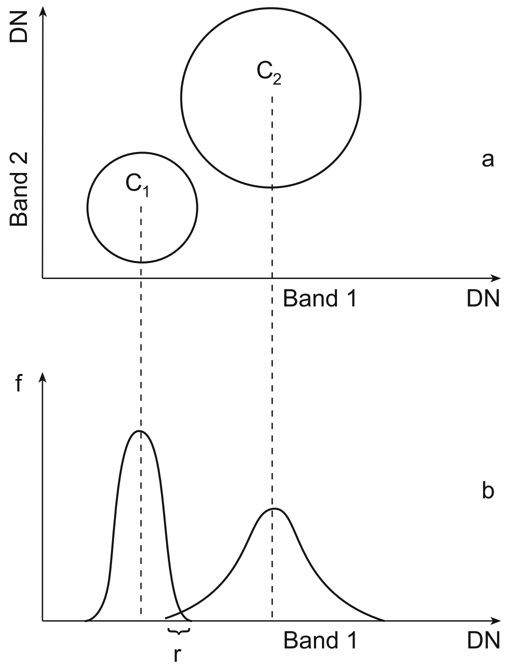


Fig. 8.72 Representation of the Gaussian probability distribution in a 2D multispectral space of two spectral classes C_1 and C_2 (**a**). In the separation histogram (**b**) r is the overlap or indecision region of data belonging to the classes C_1 or C_2. f: Digital Number (DN) frequency

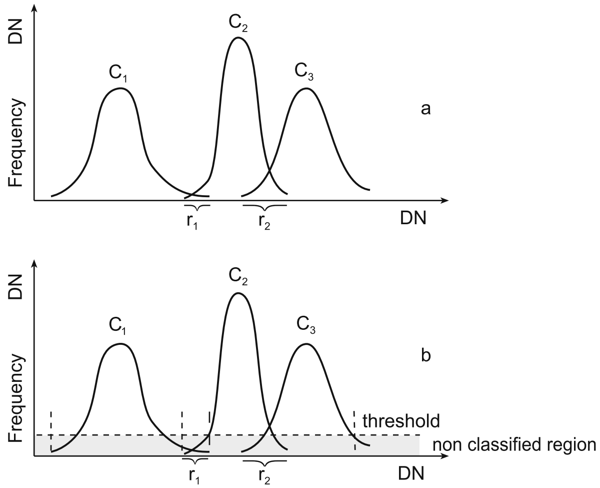


Fig. 8.73 r_1 and r_2 show the data overlap regions, and consequently indecision regions, of the Gaussian belonging distribution of two contiguous classes (C_1–C_2 and C_2–C_3). Defining a threshold is possible to reduce the indecision area labelling as not classified the pixels not included in the classification process

MLL is a hard classifier based on statistical parameters.
In this method, each thematic class is defined by a Gaussian normal distribution giving the characteristic organization of the multispectral space.

The supervised component in this classification methodology refers to the user-defined training classes. It is important that these classes are a homogenous sample of the respective class, but at the same time includes the range of variability for that class. Thus more than one training area is used to represent a particular class.

If there is considerable within-class variability, the selection of training sites can be laborious, and it is impossible to be entirely certain that a comprehensive set of training samples for each class has been specified. In many cases, it is impossible to obtain homogenous sites. A good example of this is urban areas which contain a mixture of cover types such as bare soil/concrete, buildings, vegetation and a high degree of shadow.

For each training class, the multispectral (multi-band) pixel values are extracted and used to define a statistical signature. This signature is a statistical representation of a particular class, which is used by the decision rule to assign labels.

The a priori probability can be used to decide the weight of each thematic class in the classification.

The complexity of computation increases to the quadratic with N increasing (n¡ of spectral components or classes) and requires (N2+N) multiplications and (N2+2 N+1) additions.

Regarding the number of sample elements, the MLL requires the theoretical minimum of (N+1) pixels.

In practice the minimum is 10 N pixels per spectral class for the sampling, while 100 N or more are expected for a better classification.

It is a probabilistic model and gives information about data distribution in n spectral dimensions. The pixels are assigned to a class or another according to the belonging maximum probability criterion; moreover it is possible to define a threshold to fix the decision boundaries.

Minimum Distance

The *Minimum Distance* (MD) algorithm is a parametric supervised and hard classification method where the data distribution direction is not considered because only the distances between the pixels means and not the covariance matrices are calculated. The a priori probability is not taken into account by the algorithm and decision limits are linear.

As far as complexity of computation is concerned, this method is quicker than the MLL, less complicated and requiring N multiplications and N addictions per pixel. The complexity linearly increases with N increasing, where N is the number of classes or spectral components, and therefore the selection of another number of spectral components is useful.

The required sample pixels number can be limited, as only their average is calculated. It is admitted that the model is asymmetric in the spectral space, thus it does not give information about how the data are distributed in the *n* spectral planes.

The minimum distance is a probabilistic model where each pixel is assigned to the class with the mean value nearest to the coordinates of the considered point and which can schedule the definition of threshold levels to impose the decision boundaries.

Box Classification

In comparison with the two previous methods (Maximum Likelihood and Minimum Distance), the algorithm of the parallelepiped defines the regions in the multispectral space with the following limits:

- the a priori probability is not considered by the classifier;
- correlated data produce a spatial overlap between two or more parallelepipeds;
- separation limits among the classes are linear.

Regarding the complexity of computation, this method is the quickest and the easiest. Simple controls are carried out to assign the pixels to the drawn parallelepipeds.

With parallelepiped classifier, the whole bi- or multi-dimensional space has to be covered in order not to exclude pixels from the classification. The model is based on the choice of thematic classes on the histograms and scatterograms of the *n* bands assigning then each pixel to the correspondent parallelepiped defining a class (Fig. 8.74).

The algorithm does not provide information on the data in the *n* spectral planes and a threshold to fix the lower and upper limits of each class can be defined.

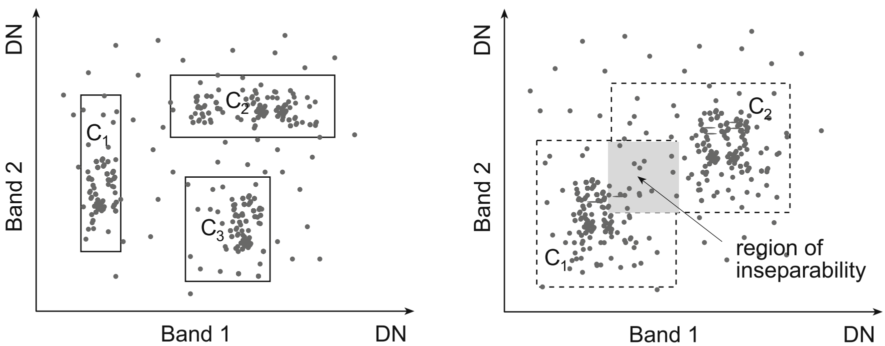


Fig. 8.74 Parallelepiped classification in three classes in a 2D space: the pixels not included in C_1, C_2 and C_3 are not classified

Fuzzy Sets

Fuzzy Sets Theory is a valid mathematical instrument for processing the information uncertainty in remotely sensed data. In *fuzzy* representation, the classes are defined by *fuzzy sets* and the pixels as elements which can partially belong to many fuzzy sets at the same time. Hence contrary to the traditional classification approaches based on the Classic Set Theory (classic is the example of the Maximum Likelihood classifier), for which is valid the law *one pixel-one class*, the pixel x of the image can be associated with a real value $f_C(x)$, ranging between 0 and 1, representing the pixel's *partial belonging degree* to the generic class C. The more $f_C(x)$ is equal to 1, the more the pixel belongs to C. This theory suits the mix condition processing, as the spectral properties of a mixed pixel located in the space among different clusters can be described in terms of partial belonging to many classes.

Statistical fuzzy classification method represents one of the theory implementations to the case of remotely sensed data classification. It defines the belonging functions f, characteristic of each class, by a statistical model based on the multivariate normal distribution, where the traditional mean and variance–covariance matrix are replaced by the *fuzzy* mean and the *fuzzy* variance–covariance matrix.

Neural Network

The interest on neural networks for multi-sensor image classification is recent; they are typically constituted by groups of elementary units (neurons) completely or partially interconnected. Network architecture depends on the considered neural pattern and can vary according to the analysed applicative problem.

An *Artificial Neural Network* (ANN) is a mathematical model based on biological neural networks. It consists of an interconnected group of artificial neurons and processes information using a connection approach to computation. In most cases, an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase.

The advantages in the use of this method are summarized in:

- a priori data statistical distribution is not needed;
- quick classification times;
- robustness to the errors.

Using neural networks, hypotheses to assign data to a given class are not used. There is no general criterion to define networks appropriate architecture, and their functioning is hard to interpret. Among the used ANN, the most important are

- *Multi-layer perceptron* (MLP) is a hierarchical structure of several perceptrons that learns non-linear function mappings. The multi-layer perceptron is capable of learning a rich variety of non-linear decision surfaces. MLPs are the most used networks for remotely sensed images classification, generally made up by three or four neurons levels, or layers: an input one, an output one and one or two hidden.

- *Structured Neural Network* (SNN), suitable to solve multi-sensor classification problems. It generates a simplified representation network allowing a detailed qualitative and quantitative representation of the particular network's way of operating. This way information about the importance of the different sensors and about the different textures used to solve the considered classification problem is acquired.
- *Probabilistic Neural Network* (PNN), pattern based on non-parametric probability assessment techniques such as Parzen (1962) algorithm. The *Parzen windows* method synthesizes an estimate of a probability density function by superimposition of a number of windows, replicas of a function (often the Gaussian).

Change Detection

Land cover/land use change monitoring by the automatic analysis of remotely sensed images acquired over the same area in different times (*Change Detection*) has a basic role in remote sensing.

The possible applications include land cover change monitoring, crop moving, deforestation, pollution monitoring, quantification of areas destroyed by fires, etc. Change detection can be performed both comparing the spectral behaviours in the multi-temporal images and comparing among them the classification maps obtained each date. In the first case, changes can be highlighted, in the second one transition between different canopies or land use can be identified. Among the first group of methods, Bruzzone and Serpico (1997) identify the following ones:

- *Univariate Image Differencing* (UID), subtracts, at pixels level, for a single spectral band, the images of a same area acquired in two different times producing an image of the differences in the two periods; the obtained *difference image* threshold recognizes changes occurred between the two acquisitions;
- *Vegetation Index Differencing* (VID), operates as the previous one, using Vegetation Indices (VI) or Tasseled Cap Transformation (Fig. 8.65) in place of the single spectral bands;
- *Change Vector Analysis* (CVA), widely used, is a robust approach for detecting and characterizing radiometric change in multispectral remote sensing data sets. CVA represents the pixels vectors in each period, operating then their difference.

The subtraction operation can be replaced by the ratio between pixels. Typically these techniques do not identify in a scientific way land cover/land use transition classes, but only the areas where the change has occurred. Only the CVA, analysing the obtained *difference vectors* phase, separates different types of change, but, as it is an unsupervised method, it does not explicitly identify the transition typologies. These algorithms are suitable to detect only one aspect: fire areas, pollution, deforestation, but cannot be used to distinguish multiple targets changes.

For the identification of classes, *the supervised classification* techniques of multi-temporal images are the most suitable. The simplest procedure is the post-classification comparison that consists in the separated classification of two

temporal images and the following verification of changes occurred in each class. A more developed technique obtaining better results is based on multi-temporal images joined classification. This technique has the advantage of using the temporal correlation in images of the same area acquired in different moments in order to improve the accuracy during the classification.

Hybrid Classification

Hybrid classification is a mixed method using both the supervised and unsupervised classifications. This approach consist in merging the two classic methodologies using at the best the characteristics and compensating the lacks of one with the advantages of the other (Fig. 8.75).

Therefore the following procedure is carried out:

- selection of areas within one or more windows in order to represent the ground spectral variability;
- application of clustering algorithms of the non-supervised classification methodology independently on each selected area;

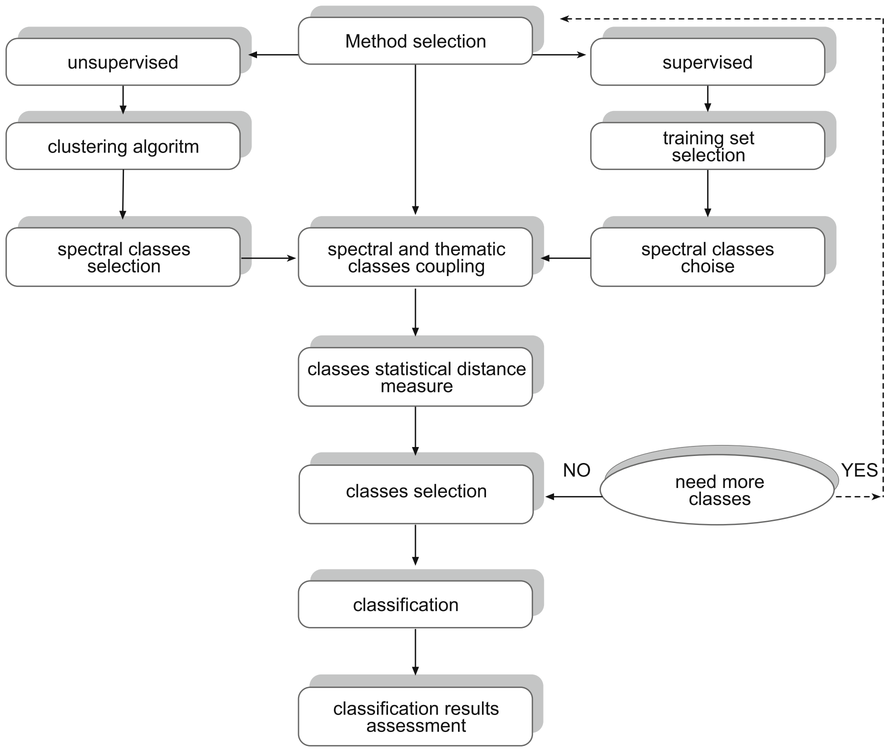


Fig. 8.75 The hybrid classification procedure

- association of the spectral classes defined by the clustering algorithm with the thematic classes corresponding to the ground truth pre-verified by field survey, study of aerial photos and available cartography;
- grouping one or more spectral classes in thematic classes;
- classification of the chosen window according to maximum probability statistical method (Maximum Likelihood, MLL), using the statistical parameters of one of the selected areas;
- classification results assessment by association of the thematic classes with the spectral classes.

8.4.3 Qualitative and Quantitative Analyses of Radar Images

Radar remote sensors measure microwave energy backscattered by natural surfaces. This scattered energy depends on the geometrical and the dielectric properties of surfaces. In vegetated areas, the volume of the canopy has a relevant impact on the backscattered signal. In the case of bare soil surfaces, the dielectric properties are directly related to the soil water content, so theoretically, radar remote sensing allows the extraction of spatially distributed soil moisture information. However, the influence of surface roughness on the scattering process limits the ability to correctly estimate volumetric values unless detailed roughness measurements are acquired.

A fundamental aspect in radar image interpretation is the so-called *speckle*. Speckle appears as a grainy *salt and pepper* texture in an image. This is caused by random constructive and destructive interference from the multiple scattering returns that will occur within each pixel. As an example, a homogeneous target, such as a vegetated field, without the effects of speckle would generally result in light-toned pixel values on an image. However, reflections from the individual plants results in some image pixels being brighter and some being darker than the average tone, such that the field appears speckled.

SAR interferometry has wide potential application in exploring delineation and density mapping of forested areas, delineation of surface water extent under adverse weather conditions, which is useful during flood mapping; detection of human settlement and crop-height estimation. This has been achieved by exploiting interferometric coherence, which is inversely related to the magnitude of random dislocation of scatterers between the two passes. These aspects are presented in Chapter 4.

8.4.4 Crop Backscattered Energy

The active (emitted by the radar system) radiation backscattered by the volume is the main phenomenon in the plant canopy behaviour. The radar waves penetrate in the plant biomass and interact with the numerous elements of the vegetation like leaves, branches, stems. Scatterers distribution, dimension and orientation determine the behaviour intensity.

In arid or semi-arid zones, this component is negligible, while at medium latitudes the backscattering due to the volume influences the signal intensity in function of the biomass of cultivated and wooded areas. Where the vegetation is composed of a higher number of layers, this scattering can contribute up to 40% of the signal; in these conditions, both the volume scattering and the geometry contribute to the signal formation.

The backscattered radiation, or *backscatter*, of a crop changes with the seasons, from the first growing phases to the ripening and to the harvesting. The variations can be explained in terms of surface geometry differences, roughness and water content.

The knowledge of the vegetation phenological cycles is at the base of the possibility to distinguish among the different crops.

In temperate zone, autumn–winter cereals crops backscatter profiles can be differentiated from other cultivated crops (Fig. 8.76).

Profiles of wheat and barley show a strong correlation with the phonological cycles; the main characteristics are the following:

- backscatter decrease in germination and biomass development phases (season beginning);
- period in which the signal has minimum values corresponding to the earing and flowering phases;
- backscatter increase in the ripening phase up to the harvesting, increasing again after the harvesting.

The minimum backscatter signal related to the earing phase is the cereals crops main characteristic. As barley ripens earlier, barley's backscatter behaviour anticipates wheat allowing in distinguishing the two crops.

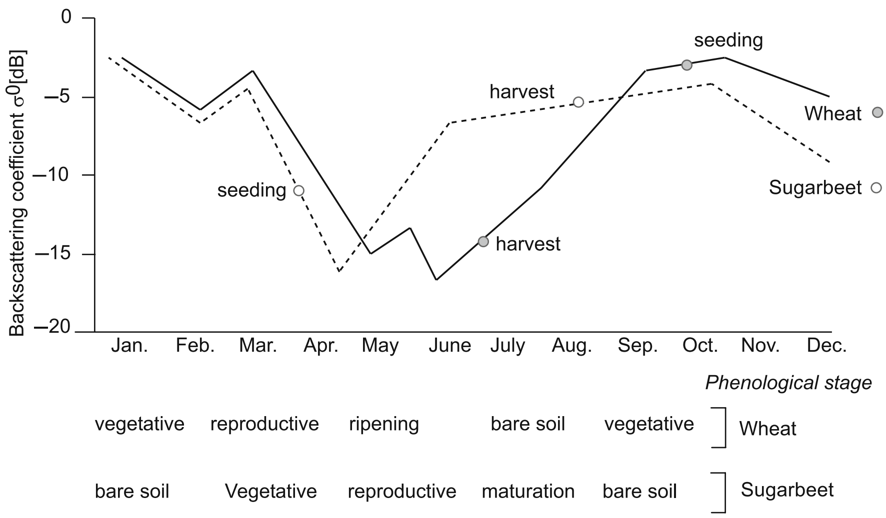


Fig. 8.76 Wheat and sugar beet radiation backscattering curves *(Bourgeaud,ESA/ESTEC, 1995)*

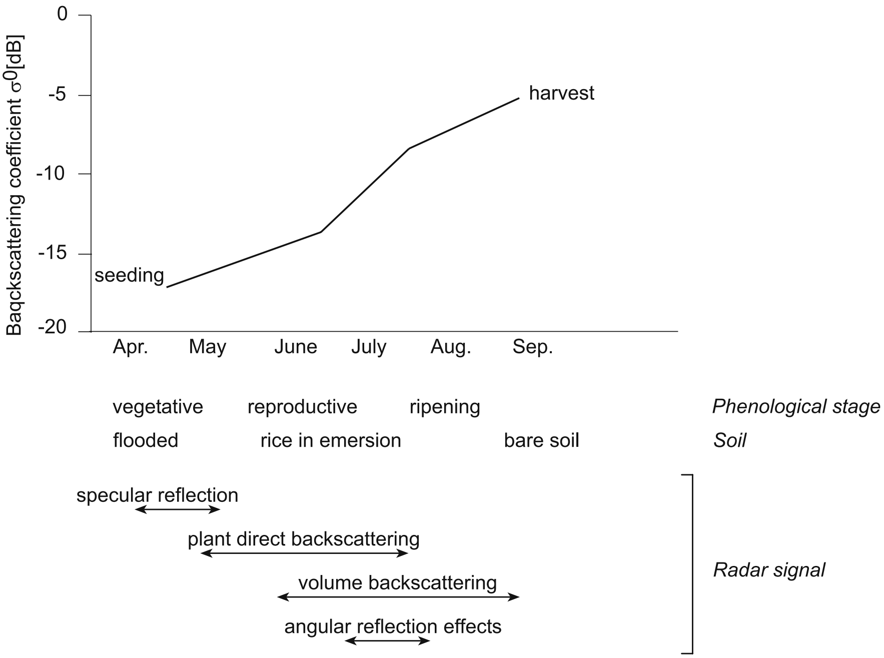


Fig. 8.77 Rice temporal signature: backscattering radiation in radar band in different cultural phases *(Aschbacher, JRC, 1995)*

Radar backscattered energy values of rice show a characteristic signature during the growing season (Fig. 8.77). This is associated to strong changes in the crop type and in the fields background in the growing-maturation phases (rice field flooding).

After the sowing, rice fields are flooded until the emerging phase. Rice biomass quickly covers the standing in the first growth phase. Paddies water retention quickly decreases during the ripening phase when the rice fields are drained. Structural roughness and moisture changes strongly influencing the radar signal.

Rice fields present a low signal (associated to darker grey tones) during the first flooding phase and in the early growing phases when the water layer is shallow; the backscatter increases up to a maximum value at the earing phase, and slowly decreases again during the senescence and the harvesting phases.

The recorded maximum value can be due to multiple reflections between vertical structure and water surface in the growing phase when signal penetration is still possible. During the ripening phase, the scattering due to the biomass volume increases, but the penetration to the water surface decreases, inducing a backscattered signal reduction. After the harvesting, the backscattered radiation decreases according to the surface roughness and moisture characteristics.

8.4.5 Soil and Water Backscattered Energy

As fields are tilled, the increase in soil surface roughness results in an increase in backscatter. Fields that are covered with significant post-harvest crop residue also experience an increase in backscatter.

If the surface has a low roughness, lower than half the wavelength of the emitted signal, the scattering is specular and the antenna does not receive any answer. Increasing the roughness the received signal intensity increases, too. Scatterer elements between half and twice the size of the wavelength give answers detectable by the antenna. The radar images appear with some degree of the radar *speckle*.

Radar penetration depends on the surface dielectric constant and is related to its porosity and to its water content.

Radar penetration in the soil is inversely proportional to the dielectric constant (E):

$$\begin{aligned} E &= 3\text{–}8 \qquad \text{dry soil} \\ E &= 80 \qquad \text{water.} \end{aligned}$$

Increasing the water content in the soil the backscattered signal intensity decreases. Moreover the penetration is directly proportional to the wavelength: the higher the λ, the higher the penetration.

Radar image pixels: each pixel in a radar image has an intensity determined by the vector sum of *scatterers*.

The back signal is dominated by:

- large-scale geometry;
- mass dielectric constant;
- small scale surface geometry (roughness).

This last aspect becomes important for roughness sizes between half and twice the radar signal wavelength. Thus for ERS-1's C-SAR band (5.6 cm) roughness values of interest vary between 2.8 and 11.2 cm.

Effects on the crops due to meteorological events (wind, precipitation) can create interferences in the definition of multi-temporal profiles. An example is cereals lodging phenomena with consequent change of the field pattern.

8.4.6 Radar Images Classification

8.4.6.1 Pixel-Based Approach

Automatic crops classification of radar images is complicated by speckle, acting as noise signals. The noise level is defined as the standard deviation of the values of pixels included in a uniform land cover area. Noise can be reduced by digital filters.

Table 8.3 Filters effect for reducing the noise in the signal/noise ratio for some land cover categories (Paudyal & Aschbacher, 1993)

Operator	Template (box)	Water	Bush	Urban	Stubble	Crops
Original	–	3.3	3.5	1.5	3.3	3.4
Medium	3×3	6.8	6.3	2.3	5.3	6.1
Median	3×3	6.4	5.8	2.2	4.9	5.5
Lee	5×5	9.5	4.9	1.9	6.9	8.3
Sigma	5×5	6.6	8.0	6.3	5.7	6.4
Frost	5×5	5.2	6.3	5.1	4.4	4.8
MAP	11×11	5.4	7.7	1.5	10.2	8.1

The presence of the noise makes the application of techniques based on single pixels classification (pixel-based) difficult; in these cases it is necessary to apply filters to reduce the noise before the classification process. The speckle has to be reduced as much as possible preserving the objects linear borders and the geometric resolution.

The results that can be obtained applying different filters to the same image are reported in Table 8.3. Filters efficacy can be evaluated also by simple image observation. Lee and MAP filters show the highest signal-to-noise ratio in crops classes with better results.

The speckle filtering is a basic requirement for radar images pixel-based classification. Merging multi-temporal radar images with optical multi-temporal and/or multispectral images, after their co-registration, several bands should be processed, slowing down the classification procedures. Therefore some methods can be adopted to reduce the number of bands to use, operating a transformation of original data into synthetic bands. Principal Components Analysis and the Canonical Analysis are the most used methods.

Many classifications, with in-field surveys, are realized based separately on multi-temporal and multispectral data, their merging and extraction of new synthetic bands. A preliminary classification accuracy assessment helps in understanding if results are satisfactory. A first draft of the thematic map is produced based on the classification of single pixels, and a majority filter is applied in order to simplify the cartographic representation. In-field truth survey and final classification accuracy statistical assessment give the definitive answer to the classification performance (Fig. 8.78).

8.4.6.2 Field-Based Approach

The *field-based* approach uses the signal mean value for each field (*Mean Field Backscatter*) or the statistically most represented value (*Majority Field Backscatter*) of each field to overcome the problem of the noise in radar images (speckle). The fields are defined as units that do not consider their internal variability (Plate 8.9).

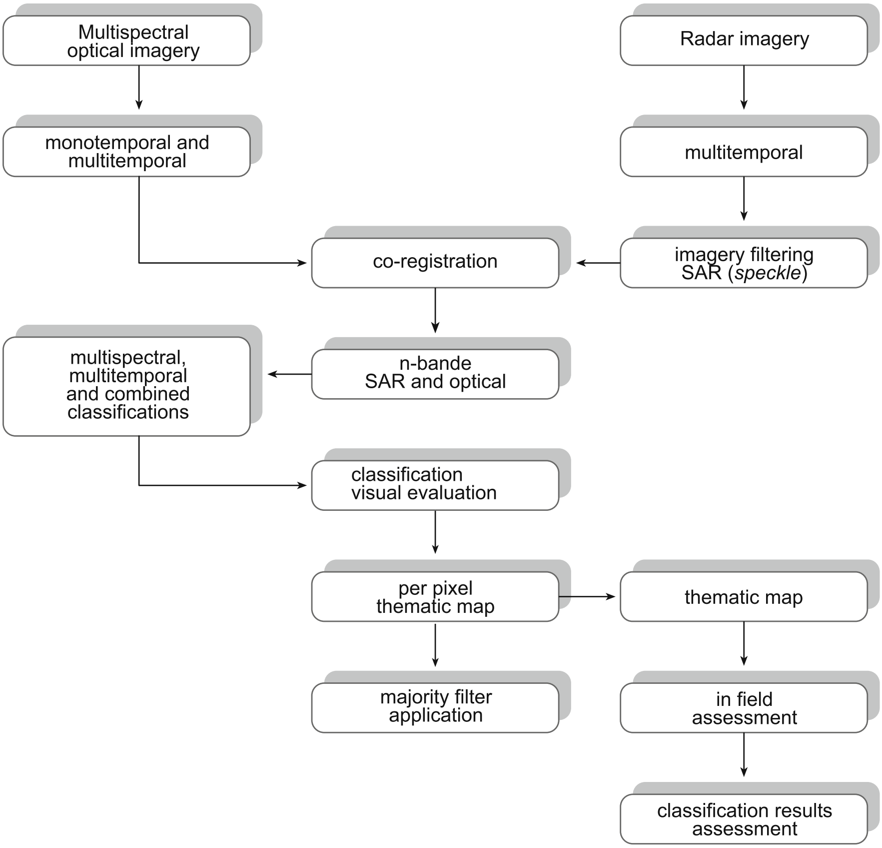


Fig. 8.78 Pixel-based classification using both multi-temporal radar and multispectral optical imagery

An example of field-based classification of integrated radar images is shown in Fig. 8.79 where the data are stored in a *Geographical Information System* (GIS). The medium backscatter value is related to the pixel contained in each field in each multi-temporal radar image. The field mean value is assigned to a crop spectral response to which a Maximum Likelihood classification is applied. The polygons can be obtained from radar images or satellite images by visual interpretation and manual digitization, or through techniques that automatically detect the field borders.

8.4.6.3 Combined Use of SAR Data and Optical Multispectral Data for Crop Classification

Synthetic Aperture Radar (SAR) multi-temporal data and visible, near and medium infrared can be efficiently integrated for crop classification.

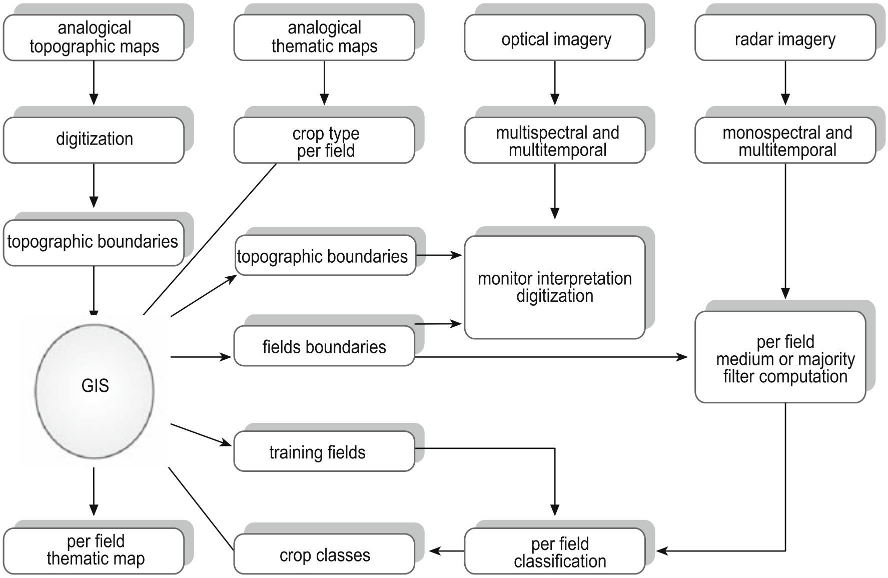


Fig. 8.79 Field-based classification using multi-temporal radar, multispectral optical imagery and ancillary data as topographic maps and surface reference data

For one single date of acquisition, the classification obtained by optical data is more accurate than from SAR images thanks to the multispectral available data.

The use of many SAR images on different dates can improve the classification accuracy from 30 to 80%. Although optical multispectral data produce the best results in terms of classification accuracy, the integration with SAR data can improve the performance.

Transformed Divergence statistical analysis and/or *Jeffries–Matusita Distance* are methods for *feature selection* that can be applied before running a classification. From a set of 4 multi-temporal SAR images and two ETM+ multispectral images (6 reflective bands each) it is possible to reduce the number of 16 bands, sparing processing time and computer memory, still having a good accuracy for the identification and classification of the objects in a scene.

SAR and optical data integration enables

- better classification accuracy for one single detection date;
- replacing of pixels covered by clouds in the optic with pixels obtained from SAR;
- better classification results using multi-temporal images.

8.4.7 Assessment of Classification Accuracy

Classifiers, both based on parametric or non-parametric methods, hard or soft, traditional or fuzzy, with or without ground surveys, lead to the realization of a thematic

map. In function of the choice of the available input data and of the classification algorithm, it is also possible to obtain very different results according to the parameters and criteria used. Rigorous control of the classifier performances and the definition of a criterion which gives assessment of quality in the different operative conditions are important, in absolute terms as well as in relation to the behaviour of other classification algorithms.

Usually the classification verification is carried out on the basis of ground truth. The classification accuracy can be experimentally measured counting both the non-classified and wrongly classified pixels. This procedure is subjected to some errors due to the *in-field* recognition, often not precise for operative difficulties, to the difficulty of correspondence between image pixels and ground pixels and, finally, to the wrong computation of mis-classified and non-classified pixels.

The real number of pixels on which the classification accuracy is calculated is limited and is performed by sample-based techniques. A traditional procedure of verification of the classification accuracy is obtained from the confusion matrix analysis. Van Genderen et al. (1978) describe the procedure adopted for a correct pixels selection for the classification and, after it, the accuracy verification.

It is typical to split the data into training sets and test sites so that training data can be used for accuracy assessment.

This way, assessment can be optimized without the problem of over fitting. However, this works only when the training set and test sites are randomly split. For instance, if there is a strong pattern found only either in the training set or test site, the optimization process will not be able to produce a model with the pattern.

The confusion matrix is a visualization tool typically used in supervised classification. Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class. One benefit of a confusion matrix is that it is easy to see if the system is confusing two classes, i.e. commonly mislabelling one as another.

The confusion matrix reports on the lines the thematic classes (A_t, B_t, C_t) chosen for the classification with a certain algorithm, and on the columns the ground truth classes (A_v, B_v, C_v). The accordance between C_t and C_v verification results is indicated by the numbers along the main diagonal: i.e. the number of pixels that both the procedures assign to the same class.

The pixels out of the diagonal represent errors that can occur:

- *omission errors*: the algorithm refuses a true hypothesis not assigning the pixels belonging to that class;
- *commission errors*: false hypotheses are accepted assigning pixels not belonging to that class.

In the matrix's last column the number of total pixels assigned to each class is reported: on the lines the pixels located by the chosen algorithm, on the columns the pixels located on the ground truth classes selected by the operator.

The same method can be used also to compare different classifiers.

Table 8.4 Numerical example of confusion matrix to assess the classification accuracy

		Surface reference classes (v)			
		A_v	B_v	C_v	Total
Thematic map classes (t)	A_t	70	4	4	78
	B_t	20	74	6	100
	C_t	10	2	82	94
Number of surface reference pixel		100	80	92	272

In the numerical example of Table 8.4, with *v: ground truth classes*; *t: map thematic classes*, considering the thematic class B:

B_t A_v (20) : number of pixels belonging to B according to *t* and belonging to A for *v*
B_t C_v (6) : number of pixels belonging to B according to *t* and belonging, according to *v*, to the class C
A_t B_v (4) : number of pixels belonging to B according to *v* and belonging to A according to *t*
B_t B_v (74) : number of pixels belonging to B according to both *t* and *v*
C_t B_v (2) : number of pixels belonging to B according to *v* and belonging to C according to *t*
(80) : sum of pixels recognized by *v* as belonging to B
Total (100) : total number of the pixels examined on the map.

The number of pixels along the matrix diagonal (70, 74, 82) represents the pixels properly classified for the classes A, B, C. The class B on the map is made up of 100 pixels (20 + 6 actually belonging respectively to the classes A and C: *commission errors*). Among the 80 ground truth pixels of the same class A 4 + 2 have been classified as classes A and C: *omission errors*. From these numbers, the percentage values of properly classified, omission and commission are retrieved, assigning the classification accuracy level. Classification accuracy is given by the sum of properly classified pixels (70 + 74 + 82 = 226) divided by the total number of pixels (272), equal to 83%.

8.5 Summary

Satellite remote sensing offers relevant opportunities for both the objects identification and knowledge concerning surface dynamic processes. Digital Image Processing refers to the process performed to enhance the image in order to facilitate the extraction of information concerning objects in imagery. Pattern recognition is mostly dependent on the skills of the expert, who also uses mathematical and statistical instruments in image interpretation, generally aimed at the classification in function of any correspondence between the objects themselves and numerical

values. Pattern recognition is processed on images represented by a matrix of pixels (Picture Element) associated with a Digital Number (DN).

The first part of this chapter describes the pre-processing operations performed on satellite or aircraft remotely sensed images and the pre-processing procedures for geometrical distortion correction.

The second part of the chapter deals with the topic of digital image processing and qualitative and quantitative interpretations.

Pre-processing includes radiometric and geometric corrections. Distortions caused by radiometric errors generate effects on the radiometric values of the image pixels, inducing a non-representative distribution of the spectral band brightness. Radiometric errors can be caused by sensors, the geometry of the system and by the atmosphere. Geometric errors can derive from the acquisition system, the atmosphere, panoramic distortion and from the Earth, as rotation and curvature.

Remotely sensed digital images can be studied by qualitative and quantitative analyses for the extraction of information. The qualitative analysis is related to human interpretation skills, the quantitative analysis to the automatic digital processing or semi-automatic interpretation, if there are controls by the operator in one or more phases of the processing. Before qualitative and quantitative analyses, spectral images can be processed improving visualization and facilitating the reading and interpretation.

Spectral analysis can be performed operating on a single pixel with punctual transformations or involving more pixels.

Further Reading

AAVV, 1997, Manual of Photographic Interpretation, Philipson W.R. (Editor in Chief), 2nd ed. American Society for Photogrammetry and Remote Sensing. Bethesdo, Maryland

Avery T.E., Berlin G.L., 1992, Fundamentals of Remote Sensing and Airphoto Interpretation, Fifth ed. Macmillan Publishing Company, New York, p. 472.

Chen C.H. (Ed.), 1996, Fuzzy Logic and Neural Network Handbook. McGraw-Hill Professional Publishing Group. New York.

Lenoble J., 1985, Radiative Transfer in Scattering and Absorbing Atmospheres: Standard Computational Procedures. A. Deepak Publishing, p. 583, ISBN 0-12-451451-0. www.deepakonline.com.

Oosterom P.V., Zlatanova S., Penninga F., Fendel, E. (Eds.), 2008, Advanced in 3D Geoinformation Systems. Springer, ISBN 978-3-540-72134-5. Berlin.

Richards J.A., Xiuping J., 1998, Remote Sensing Digital Image Analysis, An Introduction, 3rd revisited and enlarged edition. Springer. Berlin.

Bibliography

Aschbacher J., 1995a, Rice mapping and crop growth monitoring. Earth Observation Quarterly, ESA, 49: 1–3.

Aschbacher J., Pongsrihadulchai, A. et al., 1995b, Assessment of ERS-1 SAR data for rice crop mapping and monitoring. IEEE Transaction on Geoscience and Remote Sensing, 2: 2183–2185.

Avery T.E., Berlin G.L., 1985, Interpretation of Aerial Photographs. Burgess Publishing Company, Minneapolis, USA, 4th Edition.

Baret F., Guyot G., 1991, Potentials and limits of vegetation indices for LAI and APAR assessment. Remote Sensing of Environment, 35: 161–173.

Baret F., Guyot G., Major D.J., 1989, TSAVI: a vegetation index which minimizes soil brightness effects on LAI and APAR estimation. Proceedings of the 12th Canadian Symposium on Remote Sensing and IGARSS'89, Vancouver, Canada, Vol. 3, pp. 1355–1358.

Bowers S.A., Hanks R.J., 1965, Reflection of radiant energy from soil. Soil Science, 100: 130–138.

Bocchi S., Gomarasca M.A., Maggiore T., Galli A., Rossi B., 1995, Rice (Oryza sativa L.) growth analysis using radiometric indices. International Journal of Agricoltura Mediterranea, 125: 341–353, n. RAISA 900.

Brigatti M., Fasolini D., Gomarasca M.A., Pagnoni F., Rasio R., 1995, Interaction of soil information system and remote sensing for identification of land use in relevant soil mapping units. International Symposium Remote Sensing and GIS as Tools for Monitoring Soils in the Environment, 6–10 February 1995, Ouagadougou, Burkina Faso, pp. 351–361, n. RAISA 2013.

Clevers J.G.P.W., van Leeuwen H.J.C., 1996, Combined use of optical and microwave remote sensing data for crop growth monitoring. Remote Sensing of Environment, 56(1): 42–50.

Colwell B.J., 1974, Vegetation canopy reflectance. Remote Sensing of Environment, 3: 175–183.

Crist E.P., Cicone R.C., 1984, A physically-based transformation of thematic mapper data: the Tasseled Cap. IEEE Transactions, v. GE 22, 3: 256–263.

Elvidge C.D., Chen Z., 1995, Comparison of broadband and narrow-band red and near infrared vegetation indices. Remote Sensing of Environment, 54: 38–48.

ESA, European Space Agency, 1995a, Satellite Radar in Agriculture, Experience with ERS-01, SP-1185, Eds. ESTEC, Noordwijk, The Netherlands, pp. 1–71.

ESA, European Space Agency, 1995b, New View of the Earth, Scientific Achievements of ERS-1, ESA SP-1176-I, Eds. ESTEC, Noordwijk, The Netherlands, pp. 1–162.

ESA, European Space Agency, 1996, New View of the Earth, Scientific Achievements of ERS-1, ESA SP-1176-II, Eds. ESTEC, Noordwijk, The Netherlands, pp. 1–162.

Estes, J.E., Hajic E.J., Tinney L.R., (Author–Editors), 1980, Fundamentals of image analysis: analysis of visible and thermal infrared data, Chapter 24. In: Manual of Remote Sensing, 2nd ed. American Society of Photogrammetry, Falls Church, Virginia, pp. 987–1124.

European Commission, 1993, CORINE Land Cover-Technical Guide. Directorate General Environment, Luxembourg, pp. 1–136.

Gomarasca M.A., Strobelt S., 1996, Identification of Unknown Waste Sites Using MIVIS Hyperspectral Images. Proceedings 2nd International Airborne Remote Sensing Conference, San Francisco, 24–27 June 1996, Vol. III, pp. 335–342.

Guyot G., 1989, Signatures spectrales des surfaces naturelles, Collection Teledetection satellitaire n. 5, Paradigme, p. 178.

Hoffer R.M., 1978, Biological and physical considerations in applying computer-aided analysis techniques to remote sensor data. In: Remote Sensing: The Quantitative Approach, P.H. Swain and S.M. Davis (Eds.). McGraw-Hill Book Company, International Editions, New York pp. 227–289.

Huete A.R., 1988, A Soil-Adjusted Vegetation Index (SAVI). Remote Sensing of Environment, 25: 295–309.

Hulbert E.O., 1945, Optics of distilled and natural water. Journal of the Optical Society of America, 35: 698–705.

Jacobsen K., 2007, 3D remote sensing. Status report. Proceedings of the 27th EARSeL Symposium, GeoInformation in Europe, M.A. Gomarasca (Ed.). Millpress, The Netherlands, pp. 591–599, ISBN 9789059660618.

Kaufman Y.J., Tanré D., 1992, Atmospherically Resistant Vegetation Index (ARVI) for EOSMODIS. IEEE Transactions on Geoscience and Remote Sensing, 30(2): 261–270.

Kawamura M., Jayamana S., Tsujiko Y., 1996, Relation between social and environmental conditions in Colombo Sri Lanka and the Urban Index estimated by satellite remote sensing data,

The International Archives of Photogrammetry and Remote Sensing, part B7, Vienna, 12–18 July, Vol. XXXI, pp. 321–326.

Kauth R.J., Thomas G.S., 1976, The Tasseled Cap: A graphic description of the spectra-temporal development of agricultural crops as seen by Landsat. Proceedings Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, IN, USA, pp. 4B41–4B51.

Key C.H., Benson N.C., 1999. Measuring and remote sensing of burn severity. In: L.F. Neuenschwander and K.C. Ryan (Eds.). Proceedings Joint Fire Science Conference and Workshop, Vol. II, University of Idaho and International Association of Wildland Fire, Moscow, ID, p. 284.

Key C.H., 2005, Remote sensing sensitivity to fire severity and fire recovery. In: E. Cuvieco, 5th international workshop on remote sensing and GIS applications to forest fire management: Fire effects assessment. Universidad de Zaragoza, Spain pp. 29–39.

MacQueen J.B., 1967, Some methods for classification and analysis of multivariate observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley, Vol. 1, pp. 281–297.

Markham B.L., Barker J.L., 1986, Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectances and at-satellite temperatures. EOSAT Technical Notes.

Martin M.P., 1998, Cartografia e inventario de incendios forestales en la Peninsula Iberica a partir de imagines NOAA AVHRR. Doctoral thesis, Universidad de Alcalà, Alcalà de Henares.

Oosterom P.v., Zlatanova S., Penninga F., Fendel E. (Eds.), 2008, Advanced in 3D Geoinformation Systems. Springer, Berlin ISBN 978-3-540-72134-5.

Orlov D.S., Bildebaeva R.M., Sadovnikov Yu N., 1976, Quantitative laws of reflection light by soils, sub-tropical soils of Western Georgia. Scientific Papers of High School Biology, 22: 127–132.

Paine D.E., 1981, Aerial Photography and Image Interpretation for Resource Management. John Wiley and Sons, New York, p. 571.

Parzen E., 1962, On estimation of a probability density function and mode. Annals of Mathematical Statistics, 33: 1065–1076.

Pedrycz W., 1990, Fuzzy sets in pattern recognition: methodology and methods. Pattern Recognition, 23: 121–146.

Perry C.R., Lautenschlager L.F., 1984, Functional equivalence of spectral vegetation indices. Remote Sensing of Environment, 14: 169–182.

Pearson R.L., Miller L.D., 1972. Remote mapping of standing crop biomass for estimation of the productivity of the short-grass Prairie, Pawnee National Grasslands, Colorado. In: Proceedings of the 8th International Symposium on Remote Sensing of Environment, ERIM, Ann Arbor, MI, pp. 1357–1381.

Pinty B., Verstrate M., 1992, GEMI: a non-linear index to monitor global vegetation from satellites. Vegetation, 101: 15–20.

Qi J., Chehbouni A., Huete A.R., Kerr Y.H., Sorooshian S., 1994, A Modified Soil Adjusted Vegetation Index (MSAVI). Remote Sensing of Environment, 48: 119–126.

Richardson A.J., Wiegand C.L., 1977, Distinguishing vegetation from soil background information. Photogrammetric Engineering and Remote Sensing, 43(2): 1541–1552.

Ridler T.W., Calvard S., 1978, Picture thresholding using an iterative selection method. IEEE Transactions on Systems, Man and Cybernetics, 8: 630.

Rondeaux G., 1995, Vegetation monitoring by remote sensing: a review of biophysical indices. Photo Interpretation, 33(3): 197–212.

Rondeaux G., Steven M., Baret F., 1996, Optimisation of Soil-Adjusted Vegetation Indices. Remote Sensing of Environment, 55(2): 95–107.

Rouse J.W., Haas R.H., Shell J.A., Deering D.W., Harlan J.C., 1974, Monitoring the vernal advancement of retrogradation of natural vegetation. Final Report, Type III, NASA/GSFC, Greenbelt, MD, p. 371.

Rouse J.W., Haas R.H. Jr., Schell J.A., Deering D.W., 1974, Monitoring vegetation systems in the great plains with ERTS. NASA SP-351, 3rd ERTS-1 Symposium, Washington, DC, pp. 309–317.

Slater P.N., Biggar S.F., Holm R.G. et al., 1987, Reflectance based methods and radiance based methods for the in-flight absolute calibration of multispectral sensors. Remote Sensing of Environment, 22: 11–37.

Slater P.N., Biggar S.F., Thome K.J., Gellman D.I., Spyak P.R., 1996, Vicarious radiometric calibrations of EOS sensors. Journal of Atmospheric and Oceanic Technology, 13: 349–359.

Smith R.C., Baker K.S., 1981, Optical properties of the clearest natural waters (200-800 nm). Applied Optics, 20: 177–184.

Stroppiana D., Brivio P.A., Zaffaroni P., Boschetti M., Mollicone D., Petrucci B., 2007, Burnt area mapping within the borders of the Italian National Parks using ASTER images. Proceedings of the 27th EARSeL Symposium, GeoInformation in Europe, M.A. Gomarasca (Ed.). Millpress, The Netherlands, pp. 211–217, ISBN 9789059660618.

Swain P.H., Davis S.M., 1978, Remote Sensing: The Quantitative Approach, Mc Graw-Hill Book Company, Paris, New York.

Thome K., Markham B., Barker J., Slater P., Biggar S., 1997, Radiometric calibration of Landsat. Photogrammetric Engineering and Remote Sensing, 63(7): 853–858.

van Genderen J.L., Lock B.F., Vass P.A., 1978, Remote sensing: statistical testing of thematic map accuracy. Remote Sensing of Environment, 7: 3–14.

Villa P., 2007, Imperviousness Indexes Performance Evaluation for Mapping Urban Areas Using Remote Sensing Data, Urban Remote Sensing Joint Event, 2007, Paris, 11–13 April, pp. 1–6, ISBN: 1-4244-0712-5.

Villa P., 2008, Urban Areas Changes Mapping Using Mid-resolution Satellite Data and Spectral Indexes: a case study in Milan, Italy, 1984–2003. IEEE Transactions on Geoscience and Remote Sensing – Special Issue on Remote Sensing of Human Settlements.

Zadeh L.A., 1973, Fuzzy Pattern Recognition. Academic Press, London.

Zilioli E., Brivio P.A., Gomarasca M.A., 1994, A correlation between the optical properties of Lacustrine waters and some indicators of eutrophication. The Science of the Total Environment. Elsevier Ed, Amsterdam, The Netherlands, Vol. 158, pp. 127–133.

Zilioli E., Bianchi R., Giardino C., Gomarasca M.A., 1996, Potential of Remote Sensing for Monitoring Interactions Between Rivers and Urban Settlements, Metropolitan Areas and Rivers, Roma, 27–31 May 1996, Vol. 1, pp. 119–127.

Chapter 9
Elements of Geographical Information Systems

The issues concerning air, water, soil and subsoil are parts of a unique system where human actions can generate complex situations that require efficient management. Powerful tools and services for proper planning, exploitation prevention and protection are needed. *Geographical Information Systems* (GIS) and their dynamical real-time relatives, the *Decision Support Systems* (DSS), are aimed at this goal: they can process a huge amount of data describing the environment in an advanced mode both during ordinary and extraordinary situations. The combined use of such tools with the *Expert Systems* (ES) allows generating scenarios useful for decision makers in land management.

GIS are informatics (software) devices that combine the semantic component of the terrestrial features (mainly described by attributes organized in tables) with their rigorous georeferenced geometrical representation (Fig. 9.1). They are equipped with many processing and representation tools that enable an effective representation of the territory as it would not be possible by other instruments; the possibility of representing and managing Digital Surface Models (DSM), for instance, permits the creation of 3D virtual worlds with a very high communication impact.

Such impressive solutions are not enough to justify and support these systems, where the basic requirement remains the sensibility of the operators during collection, processing and organization of data. Approaching the GIS without opportune skills in cartography and survey can represent a risk for a correct use of it and for the reliability of the produced results and scenarios. Nowadays a strong boost to the diffusion of the GIS comes from the Internet. The availability of distributed data and databases and the possibility of requiring and publishing them from and to remote

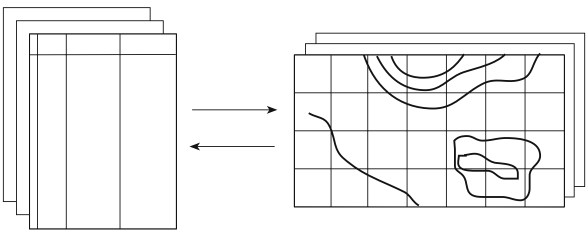


Fig. 9.1 Geographical Information Systems (GIS) combine semantic and geometrical information

M.A. Gomarasca, *Basics of Geomatics*, DOI 10.1007/978-1-4020-9014-1_9,

locations on the net are pushing the WebGIS to become a very important tool for the future complete diffusion of GIS and DSS.

Some fundamental aspects of GIS are common to some other discussed issues of this book, and they are presented in other chapters. Database Management Systems, for example, are treated in Chapter 5: Elements of Informatics, data acquisition in the chapters concerning cartography, photogrammetry, remote sensing and satellite positioning systems.

Many books have published on this topic (many are reported in the Bibliography) and just some of them can be considered exhaustive. In this chapter, even though introducing the main GIS components and data spatial analysis models and methods, attention is given to the basic definitions, to the investigation of the error sources and propagation dynamics and to the WebGIS.

9.1 Typology of the Geographical Information Systems

The quick evolution of these systems induced an increasing number of definitions. Thus, it is important to introduce a classification where non-Spatial Information Systems are disregarded. The suggestion given by Kraus (1995) is adopted.

The commonly defined Geographical Information System should actually be called *Spatial Information System* (SIS), the latter including the first one. The SIS is a Geographical (with reference to the territory) Information System (GIS), which does not give an idea of the multi-dimensional or spatial concept.

In general, according to Kraus, SIS classes relate to three categories of maps:

- *cadastral maps* and *technical maps* produced at 1:500–1:5000 scale;
- *topographic maps* that represent objects and their shape on the Earth's surface, such as rivers, roads, buildings, fields, at scales ranging from 1:2500 to 1:100,000;
- *geographical* and *thematic maps* that can be derived from different sources of data.

According to these categories, the Spatial Information Systems assume different names:

- *Land Information Systems* (LIS) if aimed at managing cadastral information in digital format;
- *Topographic Information Systems* (TIS) if they are aimed at describing natural and artificial landscapes introducing the third spatial dimension, i.e. the altitude, by means of its digital representation granted by the DSMs, DTMs and DEMs (better explained later);
- *Geographical Information Systems* (GIS), if they are aimed at describing natural and anthropic landscapes, synthesizing many thematic contents with a higher degree of generalization. According to the application it is devoted to, the information system can be better defined as environmental, hydrographical, forest, agricultural, etc.

Spatial Information Systems can be resumed into three levels with an increasing complexity of their functions:

- *1° level*: systems aimed at representation (continuous cartography); this level can be considered a sort of digital management of geographical maps able to considerably simplify the conventional use of the traditional maps, maintaining their same effectiveness;
- *2° level*: systems for the management of both feature geometry and attributes. Attributes are organized inside databases (archives of connected tables) directly or indirectly related to the geometrical territory features; these are generally grouped in independent thematic maps as ‘Roads’, ‘boundaries’, ‘rivers’, ‘cultivated areas’, etc., defined as ‘layers’, that are equipped with pertinent tables of attributes containing even very complex technical, administrative or topographic information. Such a kind of structure allows intuitive and interactive accessing of databases to perform complex operations regarding both the geometrical entities and their related attributes. The most important benefit of a GIS tool is the possibility of generating new synthetic information from different sources of data relating to different layers and tables.
- *3° level*: systems aimed at the dynamic simulation of complex processes developing on the territory (simulators); this level of GIS is an integration of the first two levels with specific procedures that can describe, rigorously or empirically, phenomena occurring over the territory and conditioning of it. Decision Support Systems (DSS) belong to this level.

Kraus finally shows that GIS is the successful result of the integration of scientific contributions related to external and pre-existing fields. He recognizes the following (Fig. 9.2):

– *Computer-Aided Design (CAD):* developed to support architectonical planning and technical design, enhancing computation as volume, weight, etc., and allowing the 3D modelling of the drawing. Many of the procedures acting over geometry of features available inside a GIS derive from CAD;
– *Database Management Systems (DBMS);*

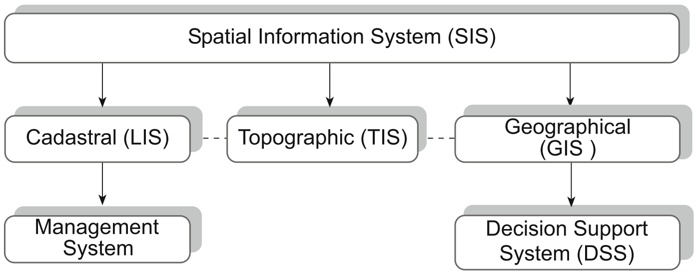


Fig. 9.2 The Geographical Information Systems (GIS) classification

– *Automated Mapping/Facilities Management (AM/FM):* they can manage geographical data both graphically and alpha-numerically. They are used mainly for technological network's management and include computational functions for spatial analysis, third dimension included. They are thought to satisfy the needs of local administrations in charge of managing technological networks for the distribution of services.

9.2 Format of the Geographical Data

A map is the most common way to represent geographical data as a set of points, lines and areas, which are positioned according to a defined coordinate system; the geographical feature, or ground element, is made up of two components:

- *positional*, which graphically and geometrically defines the position and the shape and the topology of the objects represented by geometric primitives like points, lines, polygons and pixels (e.g. country boundaries, rivers, lakes);
- *descriptive*, expressed by alpha-numerical declarations, aimed at qualifying some non-spatial properties of the geometrical features by means of attributes (numbers, strings, date); i.e. the population of each county, the toponyms, etc. Attributes are organized in tables.

According to this concept, GIS is a complex system able to manage:

- *geographical position of objects*;
- *their attributes*;
- *their spatial and attribute relationships*;
- *time factor of dynamic phenomena*.

Even if the *geographical position* of an object can be described through alpha-numerical attributes (their toponyms, the name of the administrative unit they belong to, i.e. council or province, addresses), the basic descriptors of such a property are its geographic coordinates referring to an opportune reference system (Fig. 9.3, Plate 9.3).

The *attributes* define and characterize some properties of the georeferenced-represented objects. If, for example, the object is a forest, the species composition, the height of the trees can be considered its attributes. The attributes are generally *non-spatial data*, as they do not have an intrinsic positional value; this means that they cannot change according to the variations of the map scale/projection because they do not have permanent reference to other entities (Fig. 9.4).

Four types of attributes are distinguished:

- *metrical*: numbers associated to the map elements that describe their geometrical properties as position, length, perimeter, area;

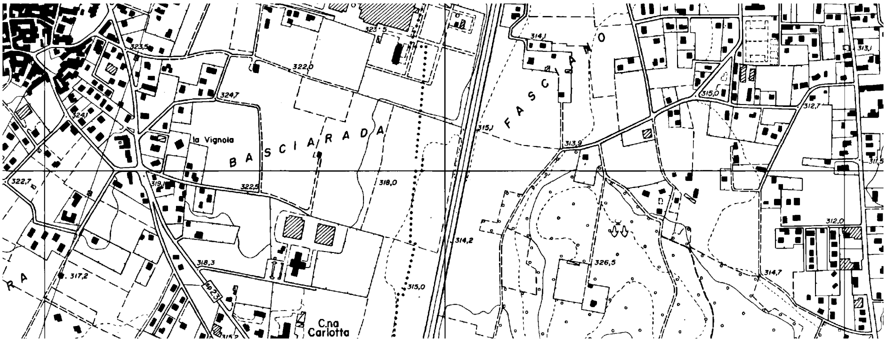
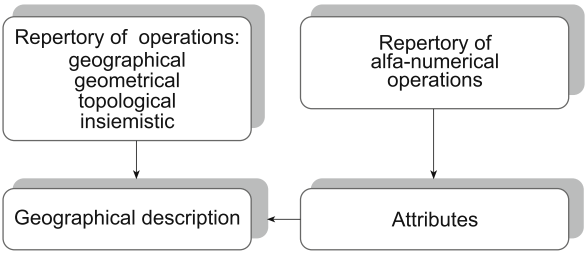


Fig. 9.3 Example of a topographic map, nominal scale 1:10. 000, where several layers (road, urban area, etc.) are shown

Fig. 9.4 Geographical data: the two components of the geographic data; a: positional, describing the topology of the objects by points, lines, polygons; b: descriptive, based on the attributes describing the properties of the geometrical features



- *graphical*: numerical codifications useful for the representation of map features, such as the symbol used to represent opportunely a point, or the colour (or the thickness) to make up a line to improve the map interpretation;
- *descriptive*: alpha-numerical declarations useful to qualify the features; for example the name of the owner of a certain building (polygon feature), its address;
- *complex*: external attributes used to describe cartographic elements, for example textual description of an industrial polygon in a land use map.

Spatial relationships, among the geographical elements, can be defined as (Fig. 9.5):

- *topological*: equivalence, partial equivalence, inclusion, adjacency, separation;
- *direction*: front, back, over, below, cardinal directions and their combinations, metrical descriptions of angles;
- *qualitative proximity*: near, far, next to;
- *quantitative*: distance measurements.

Relations can be simple or very complex. It is important, for example, not only to localize areas in danger of fire and basins for water caption by helicopters but also to know how far they are and how long it takes the helicopter to reach the place. By

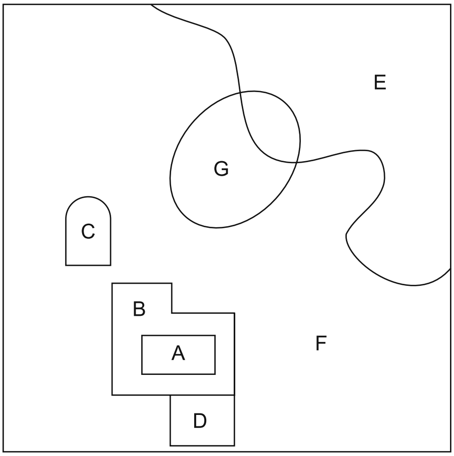


Fig. 9.5 Spatial relationships among the geographic elements
– topological: A internal to B, D connected with B, C disjointed from B, G overlapping E;
– directional: G east to C, C north to D;
– proximity: C near B, D far from E.

studying the combinations of different elements, some intersections or associations among particular phenomena can be found: for example, it is possible to look for a relation between the incidence of a disease and the presence of one or more known pollution sources.

The *time factor* allows temporal analyses through the characterization of the data in time in order to gather precious information about the evolution of a certain event, to model the phenomenon and to generate forecasting useful for early warning and early intervention.

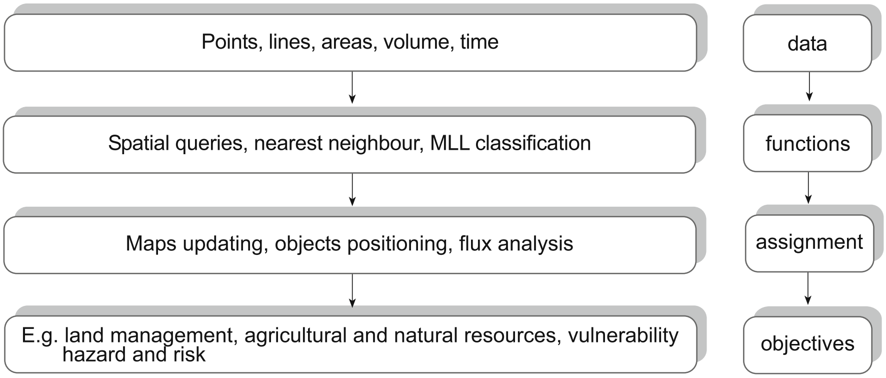


Fig. 9.6 Logical organization of a GIS

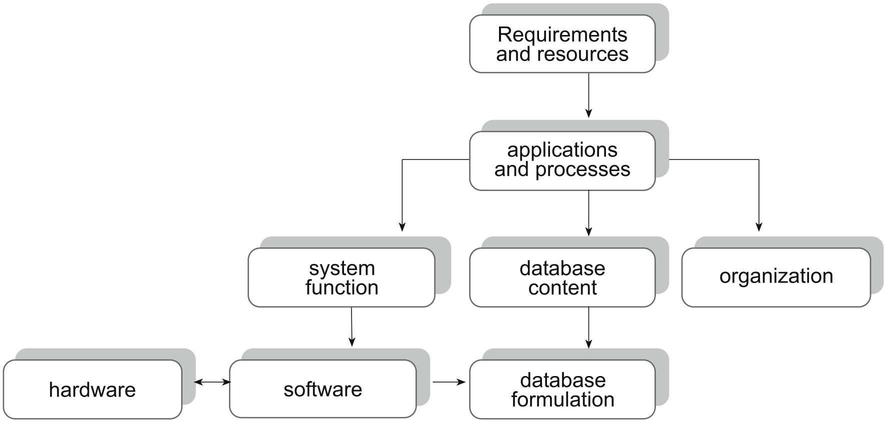


Fig. 9.7 The structural organization of a GIS: the application must address the system configuration

9.3 GIS Components and Structure

For an organic structure, it is important to specify the GIS components and its organization, as well as the processing logic that defines its structure (Fig. 9.6). These topics have already been treated in Chapter 5. GIS main components are

- hardware;
- software;
- data;
- organization context or liveware (Fig. 9.7).

9.3.1 Hardware

It includes all the components of the calculator (a Personal Computer (PC)) that hosts the system and all the peripherals. The performance of the PC strongly conditions the efficiency of the software tools and of the peripherals for data insertion and for the representation of the cartographic products, such as digitizer, scanning instruments, printers.

9.3.2 Software

Software specifically designed for geographical data processing is capable of managing vector, raster or hybrid data.

The main software components of a GIS are

- the *Database Management System* (DBMS);

- the *basic functions, managed through* an opportune user interface, are dedicated to:
  - data input and validation;
  - data storage and database management;
  - data analysis and processing;
  - data output.

The Structured Query Language (SQL) controls the basic functions. SQL database is a computer language designed for the retrieval and management of data in *Relational DataBase Management Systems* (RDBMS), database schema creation and modification and database access management.

9.3.3 Input Data

Data are essential for a GIS; the quality of every reproduced result depends on the data availability, accuracy and homogeneity. The cost of the data collection step exceeds both the software and the hardware costs, and it is estimated as 70% of the total cost needed to obtain the output final product. Thus, it is essential that each data file entering a GIS is equipped with information declaring its quality. This is called *metadata* and must communicate the origin, the reliability, the precision, the completeness, the consistency and the updating of the data it refers to. While managing a large amount of data, another fundamental factor is represented by the constant and appropriate updating of the information. To make the data available and useful to the highest number of users, they have to correspond to a pre-defined standard, or data transfer standard.

The working steps while using an already structured GIS are

- *data acquisition*: data sources can be ground surveys devoted to the collection of continuous or discontinuous punctual data, aerial and satellite images, ground and/or aerial Laser Scanning Systems (LSS), existing hardcopy maps, socio-economical databases, statistical databases, etc.;

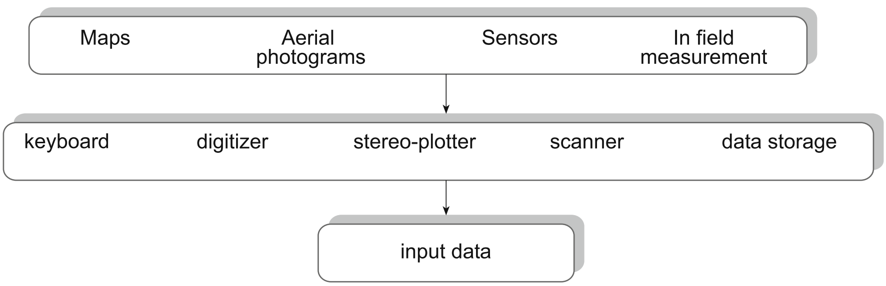


Fig. 9.8 Data collection and data input in a GIS database

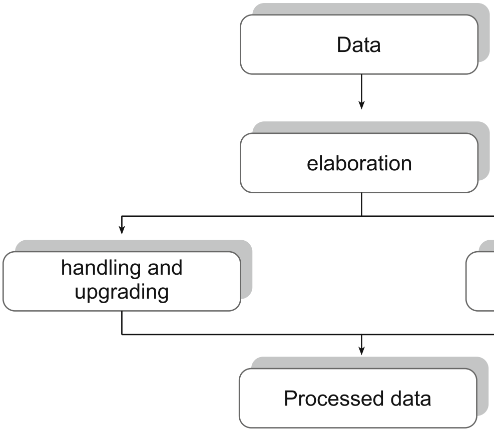


Fig. 9.9 Data transformation for their management in a database

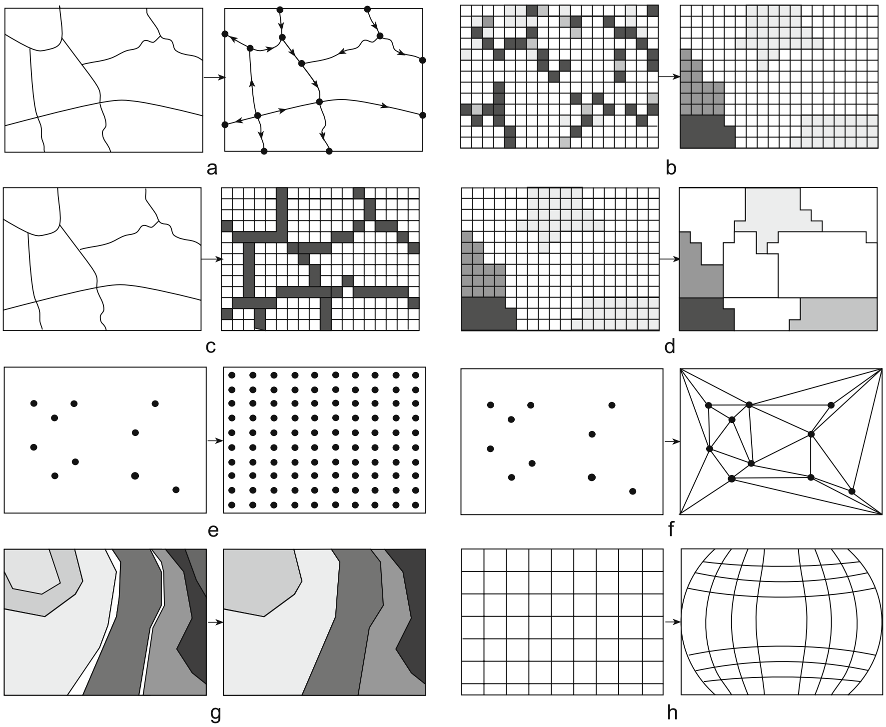


Fig. 9.10 The preliminary data processing includes several steps to structure, classify and transform the data to be ready for queries: (**a**) topological code, (**b**) image classification, (**c**) from vector to raster, (**d**) from raster to vector, (**e**) grid interpolation, (**f**) triangulation, (**g**) re-classification, (**h**) transformation of the cartographic projection

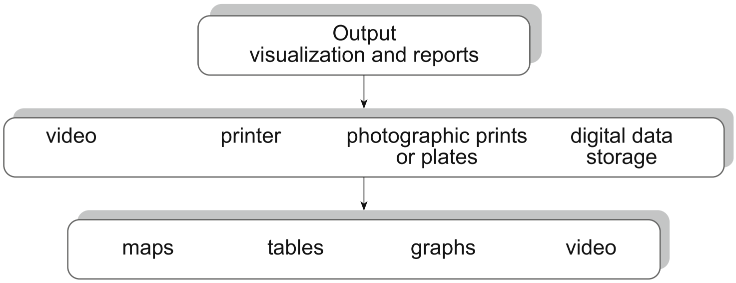


Fig. 9.11 Output data in different types of representation

- *data input*: data must be transformed into an appropriate format that can be managed by the system. The geometrical data input can come from digitization of hardcopy maps, by photogrammetric stereo-plotting, by scanning systems, etc.; the alpha-numerical data input comes from keyboard typing or table computations (Fig. 9.8);
- *pre-processing*: during this phase, the data required for the selected application are made suitable to be managed by the software. For each process, it is important to perform rigorous quality tests on both alpha-numerical and geographical data, in order to qualify the obtained results (Figs. 9.9 and 9.10);
- *data management*: the DataBase Management Systems (DBMS) are the informatics tools in charge of this task. Management regards both the graphical (spatial) and the alpha-numerical (non-spatial) data (see Chapter 5). Sometimes graphical data are separately managed from alpha-numerical ones due to their different characteristics;
- *data presentation*: results produced inside a GIS are commonly represented through layouts showing thematic maps whose degree of detail depends on the density of information, which is on the scale of the layout itself (Fig. 9.11). Thematic maps are an effective tool to integrate in a single drawing both the geometry of the features and some of their attributes.

9.4 The Organizational Context

The organizational context is represented by the operator facing the computer that hosts the GIS. This is the human component of the system as opposed to the hardware and software.

9.4.1 Databases and Structures

Geographical data components (position, attributes, spatial relationships and time factor) are generally analysed and managed using properly organized databases, defined *DataBase Management Systems* (DBMS).

Three types of DBMS can be defined:

- a single database managing both the geometric and the alpha-numerical component;
- a double database, where a dedicated database for each component exists;
- a single database only for the geometric component of data connected to other external ones containing alpha-numerical data.

Databases are described in Chapter 5, as well as the hierarchical, reticular, relational and objects oriented archiving systems structure.

9.5 Spatial Data Models

Once the analysis of the user needs and of the available resources has been executed, the next step in the designing of a GIS is the *definition of an efficient data model* able to correctly (considering a specific application) operate with respect to the attended functionalities of the system (Fig. 9.12). The spatial data modelling approach refers to two different data formats: *vector* and *grid* (*raster*) (Plate 9.1).

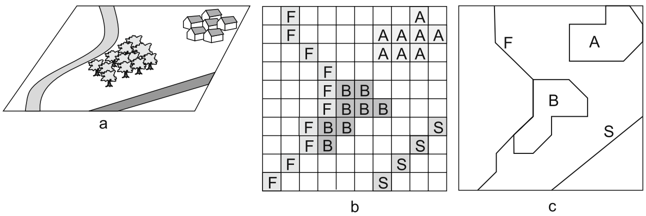


Fig. 9.12 A scene representation: (**a**) real world, (**b**) raster model, (**c**) vector model

9.5.1 Vector Format

GIS able to manage vector data appeared around the 1960s to reply to cadastral map needs and to technological networks management. In the vector data model, the territory geometry is described exploiting the elements (primitives) proper to numerical cartography (Fig. 9.13):

- *points*: which represent objects described by a pair of coordinates (i.e. a water tower, an electric power pylon, etc.);

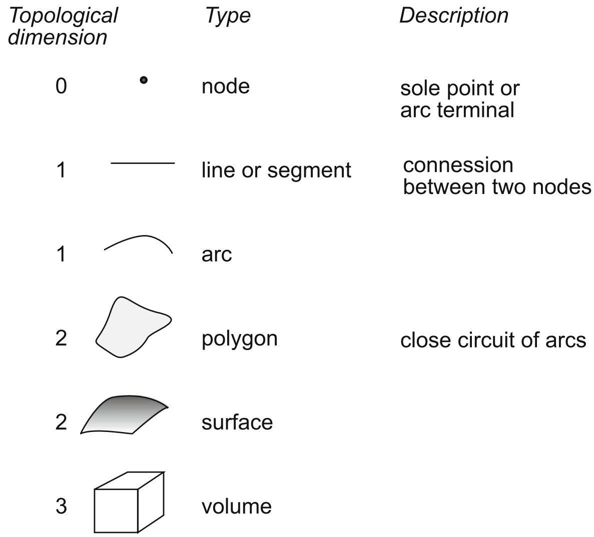


Fig. 9.13 The geometric primitives in the vector representation can be point, line, polygon, surface, volume

- *lines or segments*: represented by a set of connected points (i.e. a road); several segments, each one with two vertices, form an *arc*;
- *areas*: polygons enclosed by a polyline where the first and the last vertices are coincident (i.e. a cultivated field);
- *nodes*: particular types of points that specify a topological connection or the position of a geometric feature.

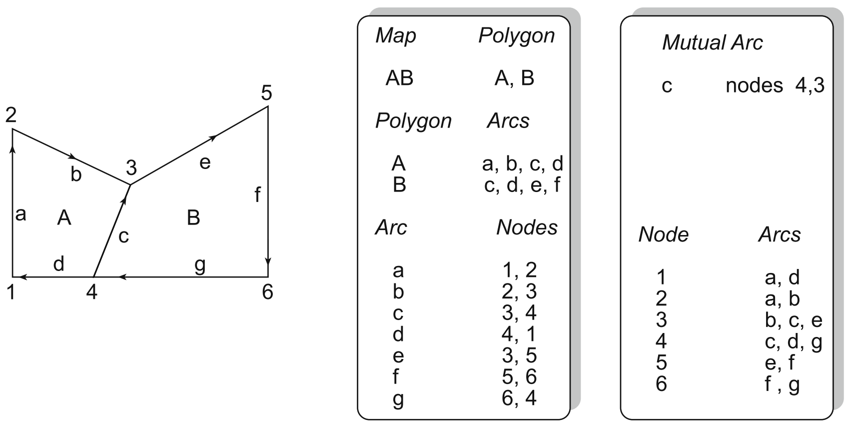


Fig. 9.14 Topological codification in a map with two polygons: each polygon is defined by arcs that are defined by nodes. Each arc has a proper direction. In the example, the arc c (nodes 3, 4) is common to the two polygons A and B

The relationships existing among the geometrical primitives (nodes, arcs and polygons) are called *topological relationships*. Their fundamental characteristic is that the arcs begin and end only in correspondence of nodes and that the polygons are enclosed by arcs. As far as an arc is concerned, its topological description must consider not only the nodes that define it but also the eventual areas (i.e. the polygons) that it is common to (at its left and right). This fact makes the arc directional, with a direction: thus the stored first coordinates of the arc refer to the starting node while the last ones refer to the ending node (Fig. 9.14). Vector GIS show relevant limits especially concerning data digitization. The constant and continuous supervision by a skilled operator is necessary, and the processing is time consuming, which often makes it impossible to have real-time updated cartographic data. This is a limitation especially for applications requiring the monitoring of evolving phenomena. On the other hand, the vector format is usually economic in terms of memory space required by the data and presents advantages in accuracy and in the quality of the representation (e.g. no degradation occurs in visualization while zooming).

9.5.2 Raster or Grid Model

GISs able to manage raster data were developed during the 1970s and were conceived as support to many subjects. In raster models, the represented space is divided into a grid, usually a regularly sized one, made by rows and columns. Each element of the grid (called cell) defines a discrete ground portion of the space. According to this data organization, the object position is defined by the row and column values inside the matrix. The single cell hosts a number defining a single attribute of the ground element in that geographic position; as only one value can be associated to each cell, if different attributes are required they must be stored in different grid files. For example, an urban area is not a defined unique polygon, but, differently from a vector representation, is a group of independent and adjacent cells whose value is a code for the attribute 'urban' (Fig. 9.15).

Even if the grid format plays an essential role in the representation of terrestrial information, in this volume the raster concept is mainly related to digital images. Burrough and McDonnell (2000) exhaustively treat this topic (Box 9.1).

The raster model presents limits especially while representing and processing independent features; in fact each object is not a stand-alone feature, but a group of cells that cannot be automatically recognized as belonging to the same shape. Furthermore, raster data require a lot of memory space to be stored. As their accuracy is strictly dependent on the ground size of the cell, which strongly conditions the file dimension, it is an important issue to look for the best compromise between the degree of spatial detail (cell size or geometric resolution) and the file size, since the latter increases as the cell's size decreases. Raster format is particularly suitable to represent thematic data, as each cell of different grid layers can refer to different attributes of that ground portion. A GIS able to manage raster data can be easily and

economically updated by digital images (aerial or satellite ortophotos) permitting the automatic re-calculation of the parameters according to the new scenario.

A particular format of the raster model based on a dynamical size of the cells is the *quadtree* model. The basic concept of this model is the increasing of the geometric resolution of the grid in the zones suffering from a high variability, thus requiring a higher detail (e.g. the borders of an area). According to this data organization if all of the cells of the grid were labelled as the same class, then the result would be a single big cell including a wide area. In a conventional raster model, a high number of cells labelled with the same attribute would have been needed for the representation of the same area. If many classes are present, the algorithm proceeds by quadrants: firstly the grid is divided into four equally sized quadrants; successively only those quadrants containing more than one class are further divided into four equally sized sub-quadrants, and so on.

This model performs some spatial analysis functions more efficiently than in the conventional raster model; the *quadtree* structure tends to have its maximum advantage in the case of management of homogeneous maps, which do not require periodical updating (Fig. 9.16).

9.6 Integration of Vector and Raster Data

GIS is strongly conditioned by the data source and the type of application. For instance a GIS planned to work with digital images is equipped with many raster processing tools, while a system conceived for managing technological network vector tools are more developed.

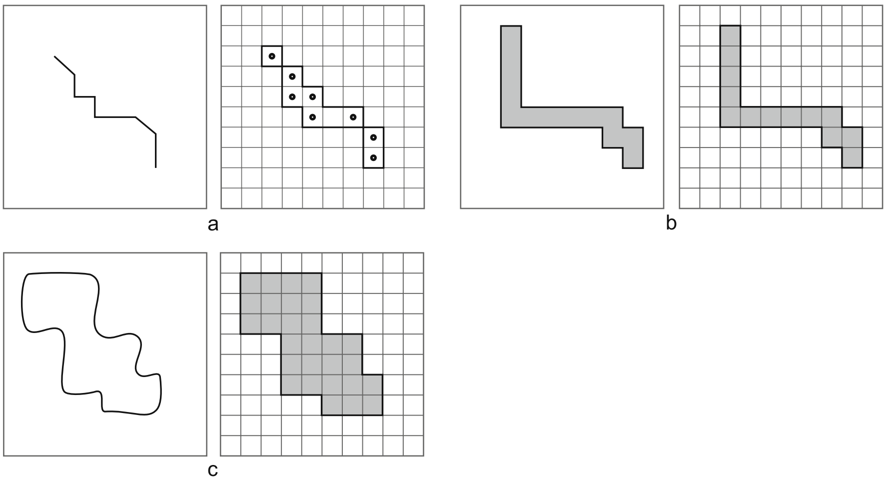


Fig. 9.15 Raster representation of terrestrial elements

Box 9.1 Strength and weakness of vector and raster models

Parameter	Vector model	Raster model
Memory occupation	Low compact data structure, accuracy and precision dependent	High data volume is resolution and compression formats dependent
Generalization	Complex	Simple intrinsic in quadtree structure
Topologic analysis	Efficient	Difficulties
Accuracy	High	Dependent from the cell resolution
Data processing	Complex	Simple
Informational layers overlap	Complex but accurate	Simple but less accurate
Definition of progressive distances (buffering)	Efficient and accurate	Simple but less accurate
Images processing	No adequate model	Efficient and accurate
DEM generation	Complex but accurate	Simple but less accurate
Data visualization	Fast	Slow and dependent from specific formats

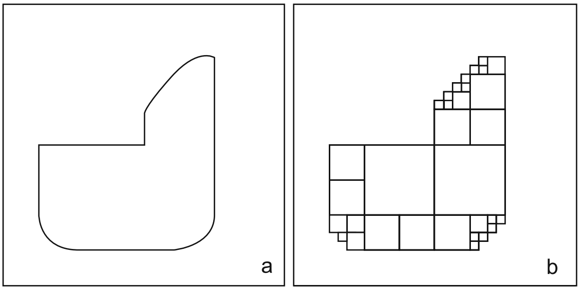


Fig. 9.16 Dynamical space fragmentation operated by the quadtree model: (**a**) polygon in a map, (**b**) representation of the polygon with the progressive use of the quadtree model in the area where more detail is required

The possibility of overlaying vector and raster data is a very useful requirement. Think about overlaying a vector 'road' layer on an aerial orthophoto: the perception of the territory is more effective, and decisions can be taken in a more complete way (Fig. 9.17). Vector data can operate as accuracy controllers of the quality of loaded orthoimages; they can be used as reference data to improve the georeferencing of raster images. On the other hand, the possibility of rasterizing vector data, trans-

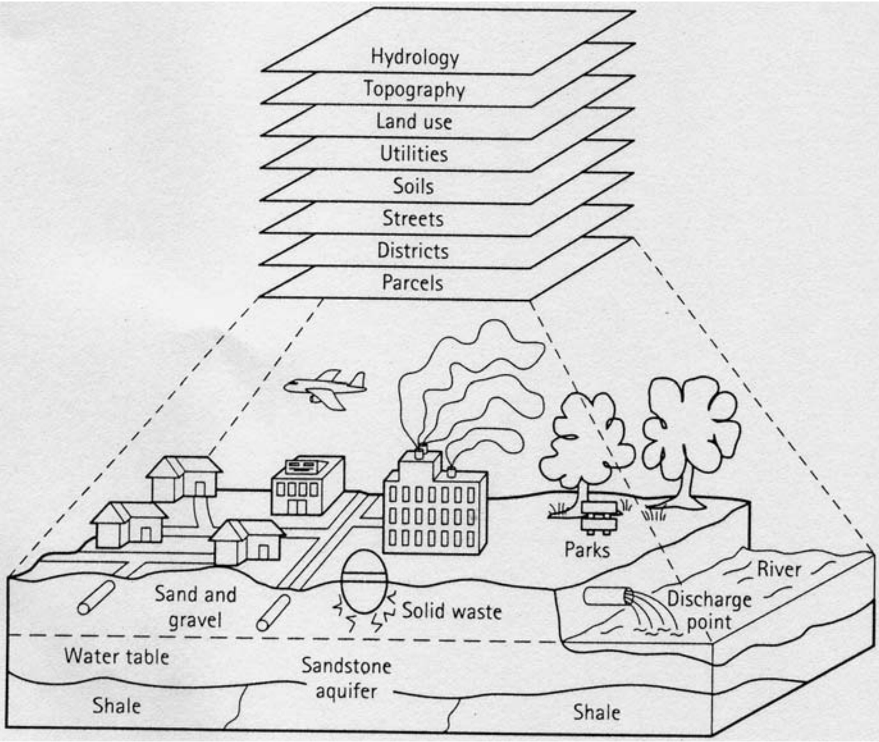


Fig. 9.17 Organization of GIS data in thematic layers. Each layer, containing specific mono-thematic information, can be combined with others to produce new layers by query (Darling & Fairbairn, 1997)

forming the representation from vector to grid, allows the processing of information through those tools (e.g. the logical operators) specifically designed for raster data and eventually to reconvert the result into a vector format.

The maximum level that can be reached by an integrated system is a complete sharing of all of the functionalities allowing a quick passage from one format to the other one according to the operation that must be performed. Some tasks are more efficient if performed in raster format, some others if performed in vector format. For example, the raster format is suitable for classification, spatial interpolation, optimization, while the vector one is more appropriate for resources allocation, for dynamic segmentation for buffering and geoprocessing.

A system where a logical integration between raster and vector data is granted can offer an easy data switch from one format to the other, a storing structure common to both the formats, an attribute management through a relational database, suitable georeferencing tools, powerful and user-friendly visualization potentialities, simultaneous raster and vector data query tools and integrated processing toolboxes.

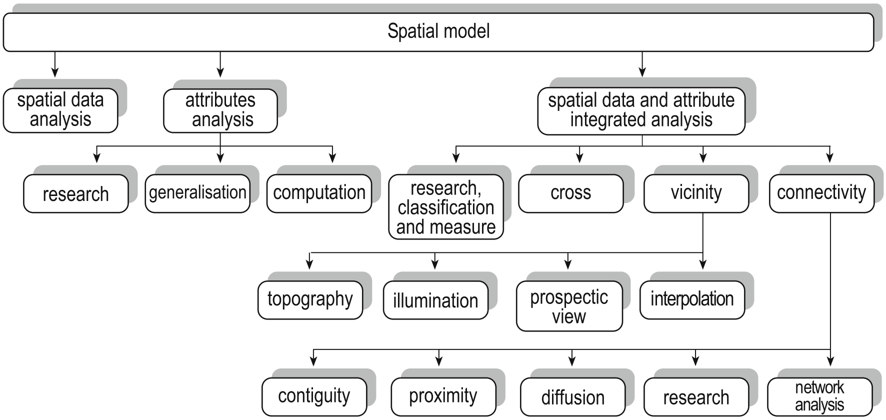


Fig. 9.18 The spatial modelling in its three components: analysis of spatial data, analysis of attributes and integrated analysis of spatial data and attributes

9.7 Methods of Spatial Data Analysis

The GIS can perform a spatial analysis; spatial relationships among the features and their attributes and the persistent link with their geometry (shape and position) make the GIS a tool able to simulate the real world and hence to help decision makers in solving actual problems and in forecasting potential consequences of risky phenomena.

The information contained in a traditional map is stored in a GIS as a sequence of thematic layers, each one containing one theme and possibly containing the same type of primitives (a polygon layer, a point layer, etc.). The *transportation* layer for example contains roads and railway lines, while the *hydrogeology* layer might include the hydrographical network, or water bodies, etc. This basic structure of data inside a GIS, where information is fragmented in layers according to the content and the type of primitives (if vector), defines the typology of operations that can be performed on the layers. For example, the length computation can be performed just on polygon (perimeter) or polyline type layers, while spatial interpolation operates on point layers, etc.

Operations can be carried out on a single data layer or by combining two or more data layers. Spatial interpolation, for example, is the most common task performed on a single layer. It is typical while generating a DEM and its derived products (slope, aspect, hill shade maps).

The development of GIS techniques has determined, in the last years, a constant increasing of the analysis toolboxes; they can be grouped in three categories: spatial data analysis, attributes analysis and integrated analysis; the last one can be further divided into 1° and 2° categories (Fig. 9.18).

9.7.1 Spatial Data Analysis

This family of operations is devoted to operate on the geometry of the layers. The same operations performed on a raster or vector model require different specific procedures. The main operations can be summarized as follows:

- data formatting in the way required by the designed GIS;
- data access (input/output);
- data georeferencing;
- data integration. Matrix operators allow merging the content of grid data. Operations among different raster layers are carried out by comparing (by formulas) the values of homologous cells belonging to each of them. Geoprocessing tools (merging, union, clipping, buffering, intersecting) are the ones concerning vector data;
- geometrical data editing by insertion of new features or reshaping of the existing one;
- map generalization by reduction of the number of the original features and their eventual transformation;
- spatial queries aimed at the selection of features satisfying topological requirements (e.g. all the hospitals within a distance of 1 km from a school; all the roads intersecting a declared administrative boundary)

9.7.2 Attributes Analysis

This group of functions is used to manage the content of the tables of attributes. Available operations are

– attribute interrogation;
– attribute editing;
– attribute analysis and reorganization;
– attribute query.

DBMS tools are those designed to perform these operations. The three main categories are

- *query functions*: they allow selective research among the records (lines) of the tables (one record per feature) looking for the occurrence of specific attribute values. Research generates new tables containing the result (a selection), without altering the source data. Basic *research* answers, for example, to the following question: 'select those features (records) having the field (attribute) *permeability* belonging to the *hydrography* table equal to X (specified value)'. Queries, that is DB interrogation, must be defined according to a well-defined syntax. The most used query language, at the moment, is SQL (*Standard Query Language*); this is independent from the database structure (Fig. 9.19);

- *generalization functions* allow feature merging based on the value of an attribute. This operation merges both the selected records of the attribute table and the geometry of the correspondent map objects. The purpose of this function is to reduce the classification level, i.e. the detail level (Fig. 9.20);
- *computation functions*: in a database, attributes are stored as numerical, string or date variables. The field of the table hosting the attribute must be defined as the correct data type with respect to the hosted attribute. According to its data type, the attributes can be processed through mathematical functions (if numerical), string functions or date functions; fields of the same type can also interact with each other to produce new results used in new fields of the table.

9.7.3 Integrated Analysis of Spatial Data and Attributes

The functions that can be used to carry out an integrated analysis, involving both the geometry and the attributes of features, can be schematically divided into four categories that can be further detailed:

- *selection, classification and measuring functions*: these operations change neither the geometry nor the attributes of features (Fig. 9.21);
  - *selection functions* perform queries typical of a GIS such as: *where are the alluvial plane units? what is the land use at this coordinate?* The result is a new layer where just the selected features are shown and equipped with the original attributes (Fig. 9.22).
  - *classification functions* permit identification and code feature categories according to a slicing of the distribution of a certain attribute value. For example, a DEM (a spread distribution of height values) can be classified (thus represented) in n height ranges (classes); a population density map can be classified according to some density ranges.

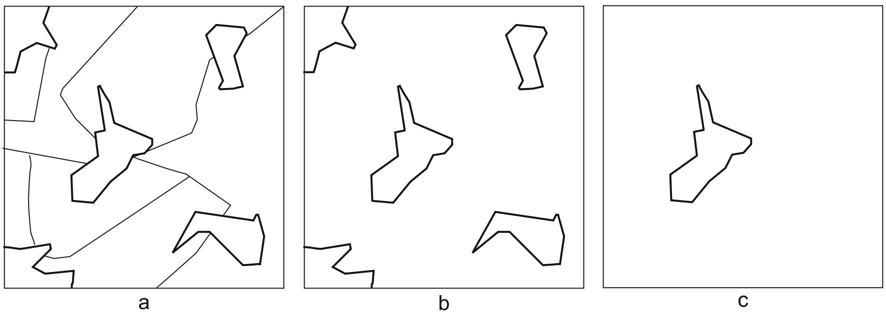


Fig. 9.19 Selection of features in a single class: (**a**) belonging to particular categories (e.g. polygon selection) and (**b**) attributes that respond to defined requirements (selection of a polygon among others)

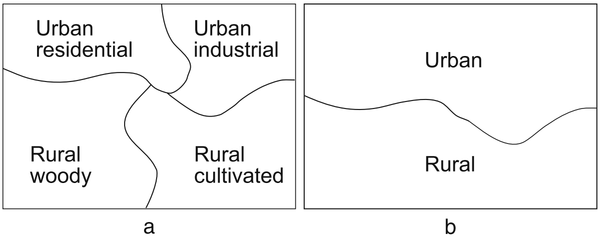


Fig. 9.20 Generalization process: (**a**) original classification and (**b**) generalized classification

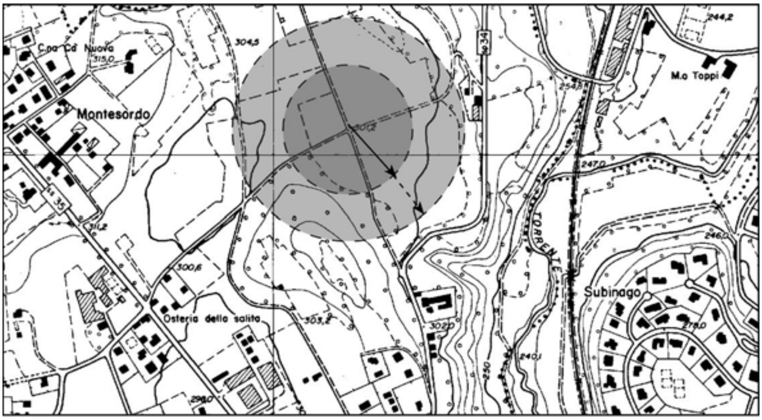


Fig. 9.21 Query of map data with geometric criterion: distance from a defined point. The result of the query could be a list of interesting entities, e.g. buildings in the test area and their emphasis on the map

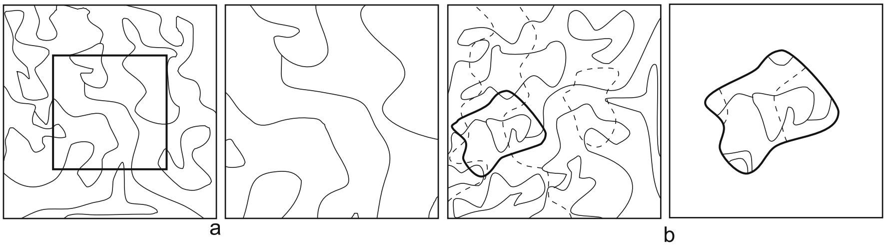


Fig. 9.22 Research in a defined area: (**a**) spatial research with a window and (**b**) spatial research based on an existing object

– *measure functions* allow calculation of lengths, areas, volumes and distances, that is to answer questions as *how long is this river? How far are the water wells from each other in a certain area? How many hectares are cultivated with wheat?*

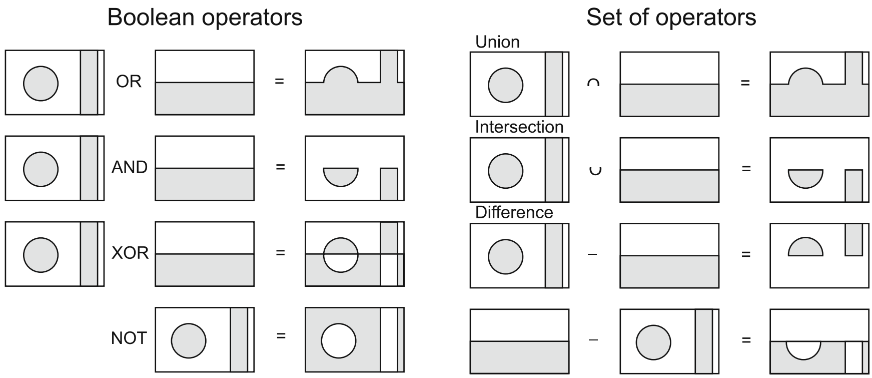


Fig. 9.23 Boolean logical operators (OR, AND, NOT) and set operators (union, intersection, difference) can be used to combine different spatial regions or layers. Database searching is based on the principles of Boolean logic, a logical relationship among research items (from the mathematician George Boole)

- *geoprocessing or crossing functions*: GIS basic functions aimed at interaction among layers. Layer are processed together to obtain new data resulting from merging, intersecting, exclusion, union, etc. operations. This does not just produce a graphical effect but generates new information at both the geometry and the attributes level. Applications are very wide, and sophisticated environmental analyses are possible through these operators. Operators useful for these tasks are called Boolean. The effect of such an approach is shown in Fig. 9.23, where a region is light (true) or dark (false) according to the satisfaction of a required condition concerning a certain phenomenon. Boolean operators execute a binary logic computation according to some specific laws that allow operation on sentences as well as on mathematical entities;

- *neighbourhood tools*: useful to evaluate the behaviour of portions of maps near a specific position (buffer zones). The four types of neighbourhood operators can be defined as follows:

  – *topographic functions*: permit the calculation and mapping slope, aspect, viewshed of a certain space-dependent function (grid format). If the function is a DEM, the results are topographical models of the area; these operations are typical of the raster model;
  – *illumination tools*: aimed at calculating the parameters related to the incoming sun radiation, hence determining how many hours a certain side of a mountain, or of a building, is exposed to the Sun;
  – *perspective view tools*: they allow visualization of the surfaces in a 3D mode improving the comprehension of some phenomena. Draping is an effective tool that easily produces 3D models by overlaying themes or digital images over a surface model (usually a DEM or a DSM, Fig. 9.24);

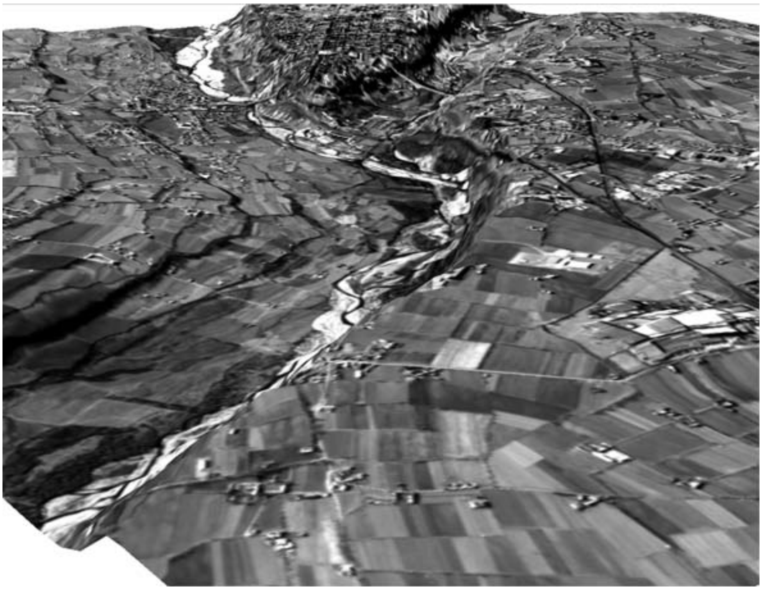


Fig. 9.24 3D modelling by perspective view of the area of the city of Cuneo, Northern Italy: the EROS panchromatic band, 1.8 m nominal resolution, is draped on the Digital Elevation Model (DEM) derived from a 1:10,000 topographic map

– *interpolation tools*: useful to over sample and regularize point distribution of an observed function. They calculate new values of the functions in new positions lying between the original widely spaced ones. Original point (or line) distribution can be either regular or irregular (Fig. 9.25);

- *connectivity tools*: useful to identify if and how segments of a network (of polygons or lines) are connected; the main ones are

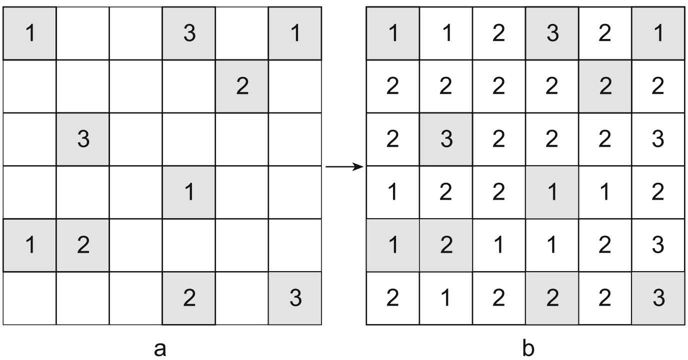


Fig. 9.25 Interpolation process: (**a**) known values and (**b**) interpolated values, in the white cells, as spatial function of the known values

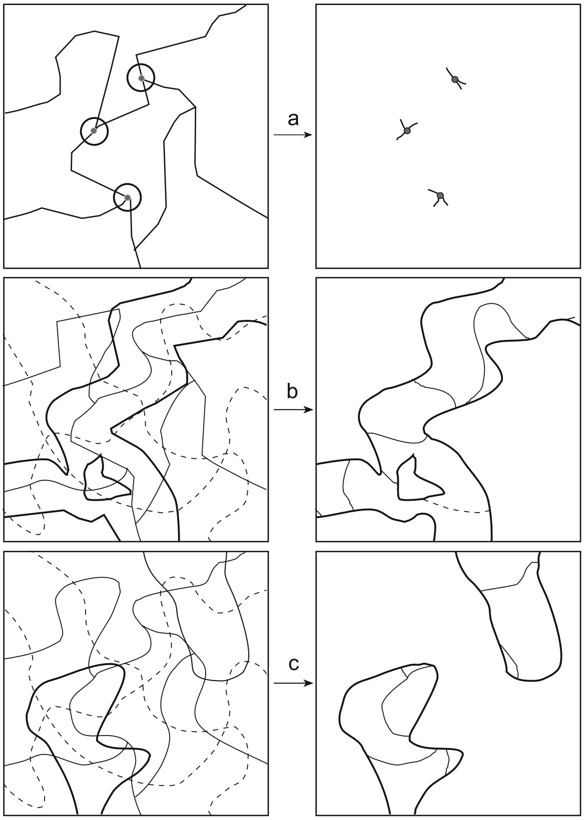


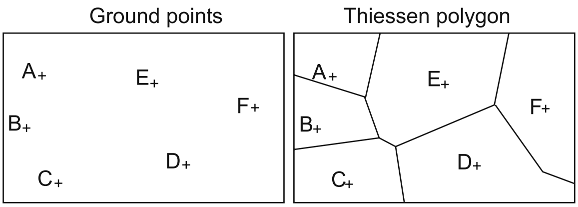
Fig. 9.26 Research based on proximity criterion. Buffering is defined specifying a distance or a radius, around a point (**a**), from a line (**b**), around an area (**c**).

– *contiguity functions*: consider those areas having common properties and evaluate the characteristics of the connected features among them. For example, to identify a suitable candidate area (sum of different polygons) as a natural reserve within a certain region, contiguity rules could be formulated. Find all the adjacent features (polygons) labelled as forest (according to a vegetation cover layer) that allow generation of a single area having at least a certain declared surface, containing enough water bodies (rivers and lakes belonging to another theme) and showing a complex morphology. The function executes these conditions considering the involved layers and proceeding polygon by polygon, or pixel by pixel, to provide an output map showing polygons, or groups of pixels, where these conditions are satisfied;
– *neighbourhood functions*: perform the computation of distances between two or more elements; the distance has to be intended as a generic cost function, in which many conditioning factors are involved. Typical examples are the determination of buffering zones around critical features (Fig. 9.26), or the generation of the Thiessen polygons (Fig. 9.27). The GIS capability to create variable and asymmetric buffering zones according to the reference mapped

features can solve complex problems and produce new thematic layers useful for decision makers;
- *spread or dispersion functions*: they investigate those phenomena whose effects over territory are related to the distance from critical features. Distance, again, is a cost function potentially depending on different constraints related to territory characteristics. For example, it is possible to calculate the dilution of a pollutant as a function of the distance from the source, from the soil type, from the land cover type, from the terrain impermeability conditions, from the slope, from the rainfall; another example is the definition of potential flooded areas with respect to the position of a dam and to the potential out coming water volumes;
- *seek functions*: used to determine the minimum path, the optimal or at least the cheapest route satisfying specific decision rules; properly applying this function on the DEM it is possible for example to define the path of water fluxes along the territory;
- *network analyses*: they are used to optimize and manage the activities considering the flow of information through network systems using research functions taking into account the way providers or users access the network and controlling the connectivity between different portions of the network. The dynamic segmentation function can associate information to any portion of a linear feature without changing its physical structure thus allowing attributes to be easily represented and efficiently managed.

9.8 Representation Methods of the Earth's Surface

In order to investigate and describe phenomena occurring on the Earth, appropriate representation models of its surface must be considered. If height measurements come from punctual and linear sampling (discrete), only an *incomplete representa-*



Points attributes

Point	Rainfall [mm]
A	3,0
B	4,5
C	2,0
D	3,5
E	-
F	-

Polygon attributes

Polygon	Area [m^2]
A	400
B	425
C	450
D	760
E	956
F	632

Fig. 9.27 Segmentation of an image, or data interpolation, using the Thiessen polygons. The examples refer to rainfall data set measured in situ

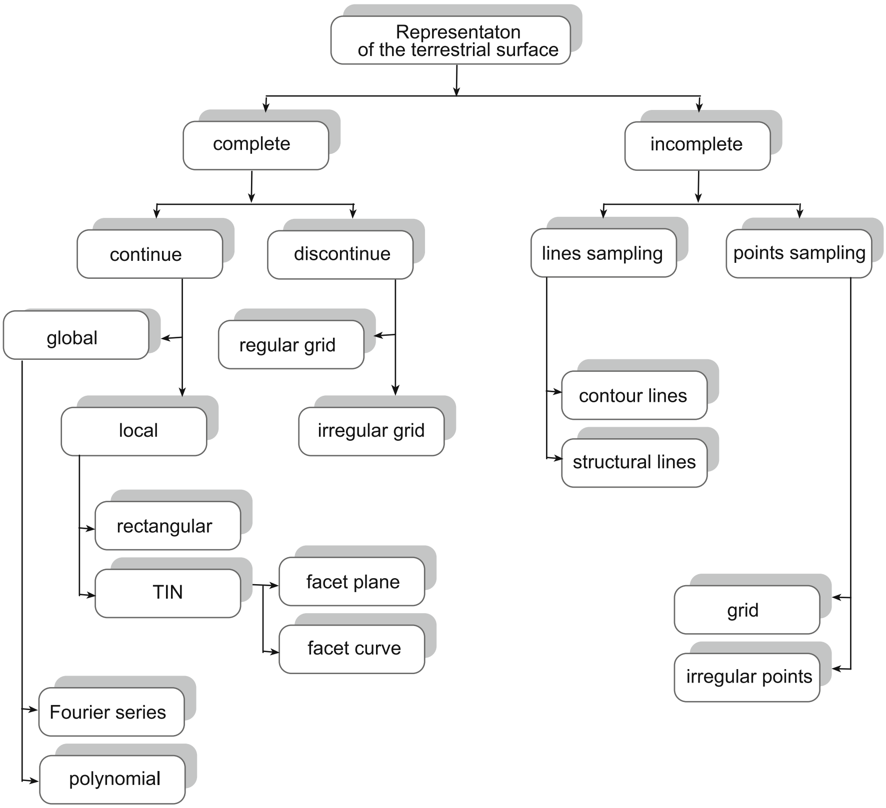


Fig. 9.28 Hierarchical subdivision of terrestrial surface representation methods

tion of the surface is possible. This often avoids the recognition of local aggregations useful for defining interesting structures and/or their variations.

On the contrary, the *complete representation* (continuous) can be seen as a series of contiguous zones, each qualified by a single code value (attribute); or it can be represented by continuous mathematical functions space dependent. It is a good rule, while describing the territory through a complete representation, to operate exploiting the aggregation of simple basic shapes to describe a more complex reality (e.g. TIN).

The hierarchical subdivision of the several representation methods of the Earth's surface is reported in Fig. 9.28; some of these are described in the following paragraph about Digital Terrain Models.

9.8.1 Digital Terrain Models

Digital Terrain Models (DTMs) are getting more and more important in geographical data processing and analysis: they allow modelling, analysis and visualization of phenomena related to the territory morphology (or to any other characteristic of the territory different from elevation).

The third dimension is a basic element of the visualization, which plays an essential role in the virtual visualization of the territory. It is important to go deeper into the meaning of the commonly used terms and understand how the Earth's surface digital models can be produced.

The digital description of ground height was born in the middle of the 20th century. J. Miller, from the famous *Massachusetts Institute of Technology* (MIT), in Boston, USA, was the first to introduce the third dimension. From that moment, the Digital Terrain Models grew in importance.Among them the DTM of the *Service Spécial des Autoroutes* (France), of the *Nordisk ADB* (Sweden) and of the MIT *Meshurface (USA)*, and the researches in the field of Linkwitz, Nakamura, Silar, Rummel in Germany, up to the most important HIFI88 method by Ebner et al. (1987) from Monaco University worked to generate. More recently, the first who approached these problems in Italy was G. Inghilleri, from Milan University, between 1964 and 1968; but already in 1833 Ignazio Porro assumed that the map is a sub-product of the topographic survey, while the ground surveys measurement accuracy should not be wasted by graphical procedures for the printing of the map. Porro wrote that the terrain should be represented by the tern of Cartesian coordinates (X,Y,Z) of all of the surveyed points; these can then be linearly interpolated: this is the basis itself of the present digital terrain models of the Earth's surface.

The classification of digital models, generally defined as the representation in digital format of a portion of terrestrial surface, is referred, in bibliography, in different ways. The most used terms are

- Digital Surface Model, DSM;
- Digital Terrain Model, DTM;
- Digital Elevation Model, DEM.

Even though some differences exist between these definitions (and relative acronyms), they are often used as synonyms. Until an international authority will unambiguously establish their rigorous meaning, they will still be used with no distinction. Nevertheless, some light differences among them can be noticed in specific publications; for instance:

- Digital Surface Model or DSM describes the terrestrial surface, including the objects covering it like buildings, vegetation and in general describes it through known geometric primitives (rectangles or triangles).
- Digital Terrain Model or DTM describes the surface just at the ground level.

- Digital Elevation Model (DEM), in geometrical terms, is equivalent to the DTM. It can be defined as a regular grid of height measurement (generally with squared cell) organized in a raster format (see Plate 9.2).

Anyway the DEM is generally used as a synonym of DTM. Afterwards and in the text, where possible, it will be generically referred to by the acronym DEM.

9.8.1.1 DSM/DTM or DEM Production Steps

The derivation or reconstruction of a 3D or 2½D from a bi-dimensional projection is called 3D scanning. To obtain the final 3D representation, it is necessary to pass through an intermediate product: a *3D cluster of points.* Nowadays it is mainly derived by the laser scanning system (Plate 9.4).

The strange concept of a 2½D surface model is related to a forcing condition in the description of the object. In fact it is assumed that for each (X, Y) position only one Z-coordinate exists. In this model X, Y planar coordinates are known data, and they are the key to access the third dimension that sound like an attribute of each point.

If the described shape is very complex, the 2½D model is not sufficient for a complete representation. For example, the 3D representation of a tree canopy cannot be visualized by a 2½D model, while a field surface can.

The problem of generating a DSM appropriate for measuring is stated as follows.

Given a group of sampled points:

$$X \in \Re^3 \tag{9.1}$$

that are assumed to belong to (or to be nearby) an unknown surface U, an approximation of it, the model S, can be created satisfying the following criterion:

$$\mathrm{d(S,U)} < \mathrm{c} \qquad \mathrm{(DSM)} \tag{9.2}$$

whered( ) : dissimilarity or distance function

c : threshold criterion.

The DSM approximation (S) is more different from reality the higher the morphological complexity of the surface U.

Nevertheless, it is important to remember that the X measurements are always affected by accidental errors and often they contain some erroneous outliers (isolated points). An example of isolated points belonging to X is an Aerial Laser Scanning (ALS) data set containing mixed measurements of both vegetated surface and bare soil. The identification and elimination of the outliers are always very difficult operations, whose efficiency can affect the final accuracy of the derived DSM, even outside their position.

The number of surface models (S) is, in practice, infinite; thus different surface models can give different good approximation of the reality. This issue makes the surface approximation problem a complex one.

A widely adopted approach to generate a DSM from the X measured points is the segmentation (regular or irregular) of the available point cloud addressed to create a *piecewise model* resulting from the integration of basic simple 3D primitives.

Different models face the problem of the control of the height differences in different ways. Some models generate a continuous geometric surface uniformly degrading in height; other ones tend to vary the data structure and the computation complexity. Each model can be efficient and accurate for the specific application it was developed for, and badly behaving in other contexts.

Thus, the quality control of a DSM depends on the use and application. A DSM aimed at producing large-scale orthophotos has to satisfy criteria different from a DSM used, for example, to realize a DEM for the study and definition of potential flooding areas at medium-small map scale. The density of the 3D point also conditions the DSM quality: larger density of points allows the model to follow the abrupt slope variations. Excessive density of points can generate higher probability of casual errors. This fact suggests selecting the data first, with a consequent loss of data frequency in the surface model.

The DSM/DEM workflow concerning both its production and its utilization develops according to the following steps (Fig. 9.29):

- data acquisition;
- model generation: definition of the model and of the interpolation algorithms;
- processing: modifications and refining;
- interpretation: information analysis and extraction from the DEM;
- visualization;
- applications.

Altimetric Data Acquisition

A 3D point survey is the most difficult and expensive step of the whole process of a DSM/DEM production. The identification of the points to survey and the sampling strategy are critical parameters for the quality of the final model. The techniques used to measure 3D points are

- ground survey;
- digitization of existing maps;
- aerial and satellite photogrammetric methods;
- radar interferometry;
- Airborne Laser Scanning (ALS).

Ground Survey allows adaptation of the sampling procedures to the ground characteristics hence providing very accurate measurements, producing very high geometric resolution DEMs. As they are expensive in terms of both money and time, they are carried out to detect the topography of very small areas, for example, contaminated sites, quarrels, landslides, volume variations.

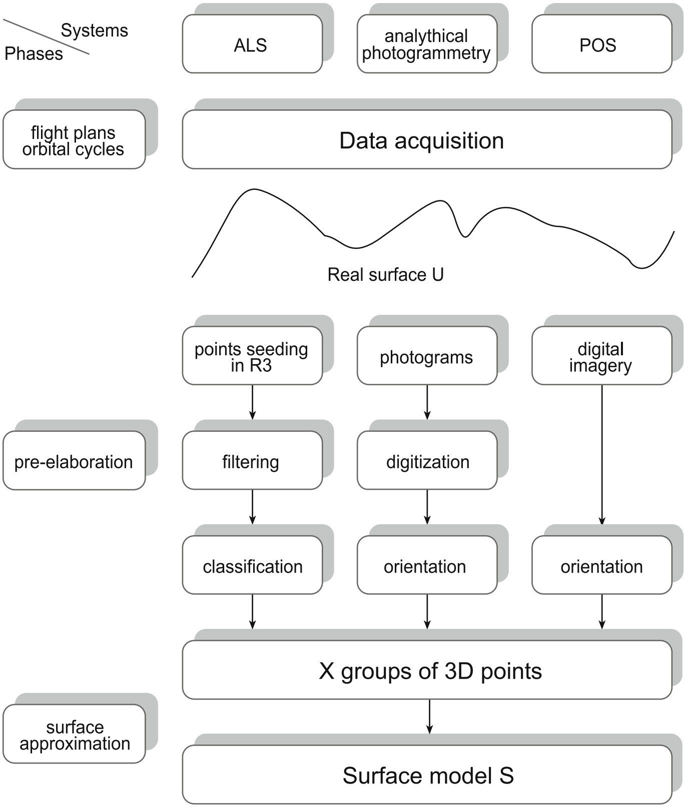


Fig. 9.29 Digital Surface Model (DSM) generation. ALS: Airborne Laser Scanning; POS: Passive Optical System

The digitization of already existing *maps* represents the data source that is most commonly used for DEM generation but which, for several reasons, is less reliable. Here are some limitations:

- the sampling density is very high along the contour lines, but there is no available information about the areas within the contour lines. For this reason during the interpolation phase, the contour lines are very often matched to points acquired in another way;
- the contour lines are digitized from traditional map products, which often are affected by uncontrollable errors;
- other defects are introduced during the digitization step.

The *photogrammetric approach* is the classic method but is surely longer and more expensive than the previous one. It allows measuring the 3D spatial coordinates of ground points recognizable over an oriented stereo-model. Wide areas can

be measured more efficiently than the ground survey. To generate a DSM, the contour lines are stereo-plotted by a skilled operator working on the stereo-model using an analytical stereo-plotter or a digital working station. To complete the job, correct eventual errors and evaluate the final quality of the obtained product, limited field surveys are required.

According to the photogrammetric approach, the height coordinate (h_p) of a point, laying on the Earth's surface, is a function of the baseline (B), of the flight height (H), of the focal length (f) and of the total parallax, as measured from the nadir position ($p_l + p_r$). As the sampling is realized in an interactive way by specialized operators, this technique can produce very accurate data sets: using standard aerial cameras having a 150 mm focal length and flying at 3000 m it is possible, in an analytical plotter, to measure elevation with an accuracy of 0.5 m. Some automatic procedures, based on autocorrelation algorithms, exist on board of digital photogrammetry working stations to extract point clouds from oriented stereo-models. These are used to process both aerial and satellite stereo-pairs. A typical example is the use of EOS-AM Aster stereoscopic satellite images that can provide, in the best conditions, Digital Terrain Models with 15 m precision.

It is worth noting that ground height measurements in a photogrammetric way are possible only in bare soil areas, but, where trees are present, ground elevation cannot be derived. In this situation, expensive ground surveys and tests by the operator are required.

Radar interferometry: medium resolution digital images acquired in Amplitude and Phase by radar systems (frequency between 500 MHz and 10 GHz) allow generating DEM. Radar systems devoted to this task are SAR (*Synthetic Aperture Radar*). In May 1991 the first European SAR was launched on board the ERS-1 satellite; a few months later the European Space Agency (ESA), after verifying the interferometric capacities, filed a patent about data processing new techniques, until it managed to demonstrate terrestrial crust movements with centimetre order precision. In 1995 the second European satellite, the ERS-2, was launched and, as suggested by the Signal Processing Unit of the *Department of Electronics and Information* (DEI) of the Polytechnic of Milan, Italy, was set on the same orbit of the ERS-1 for a short time in order to operate over the same area with 1 day delay. This setting allowed repeated acquiring of data of the entire Earth's surface, generating a DEM of wide areas. The advantages of SAR in comparison with the optical systems are related to the capacity to operate at night and with cloud cover; the SAR can provide coherent images containing measurement both of signal intensity, related to the objects' reflectivity, and of the signal phase, related to the distance between the target and the radar source.

From the interferometric phase, it is possible to obtain the relative elevation map of all SAR coordinate pixels, which then has to be put into a conventional reference system (generally UTM) by a georeferencing operation. An example of DEM of the Etna volcano generated by seven pairs of SAR images is reported in Fig. 4.39.

Airborne Laser Scanning (ALS): laser technology is now available for photogrammetric applications on board airplanes; it is a still developing and continu-

ously evolving technology that produces a huge amount of data hard to be processed, visualized and integrated into the photogrammetric process (Box 9.2).

A high-quality DSM/DEM generation requires heavy operations about the data, mainly aimed at gross error elimination, cloud filtering, simplification and segmentation. An important step is the separation between ground surface and object surfaces (trees, buildings, etc.). The approach consists in the computation of a mean surface obtained by averaging the measurements of all of the points (ground and objects). Firstly, weights are opportunely assigned in order to distinguish the points that correspond to the ground level from the above objects. In a second step, the ground points are selected and used to calculate the ground surface model. The result is strongly conditioned by the morphological characteristics of the surveyed area. For example, along the valleys, the DEM tends to model artificial depressions and to disregard the natural slopes. Thus, the final products must suffer from some manual make-ups, which sometimes also requires additional aerial/satellite images for the control (Fig. 9.30).

Box 9.2 Airborne Laser Scanning (ALS) versus analytical and digital photogrammetry, and remote sensing

Airborne laser scanning system (ALS)	Analytical photogrammetry	Digital photogrammetry and remote sensing
Active system, very powerful, collimate and monochromatic acquisition	Optical passive system	Optical passive system
Per point acquisition	Photographic emulsion	Linear or matrix CCD sensors
Polar geometry	Perspective geometry and central projection	Cylindrical geometry and central projection of the images
Direct acquisition of the 3D coordinates	Indirect acquisition of the 3D coordinates	Indirect acquisition of the 3D coordinates
No imaging, in case are monochromatic with low quality	Images with very-high geometric quality	Images with high radiometric and geometric quality
Technology and errors modelling in rapid evolution	Technology and errors modelling well known in traditional photogrammetry	Technology and errors modelling in rapid evolution
Heavy system dependence from the producer	Autonomy in photogram processing	Autonomy in image processing

9.8.1.2 3D Coordinates of the Height Spots

3D coordinates can be measured by analogical or digital collimation (manual or automatic), if an oriented image stereo-model is available. Co-linearity equations and the spatial forward intersection principle are the basic concepts of this approach. The accuracies (σ_{XY}, σ_Z) of the coordinate measurement depend on:

- quality of the orientation of the stereo-model;
- object position over images;
- flight height or object distance in aerial photogrammetry;
- orbit altitude in satellite remote sensing.

9.8.1.3 Sampling Methodologies

It is possible to define and adopt different sampling methods, according to the required task; three sampling strategies can be recognized:

- *regular*: the easiest sampling mode, in which the measured points lay down on a regular grid. This methodology, often used in a semi-automatic way, can be applied only to homogeneous and regular areas;
- *progressive*: in this case, the sampling step is adapted to the Earth's surface complexity. A higher detail is obtained in highly variable areas; the effect is a slight increase in the density of points;
- *selective*: this is the strategy adopted for all the areas characterized by a strong variability, where it is necessary to sample in a selective way along the recognized discontinuities (manual).

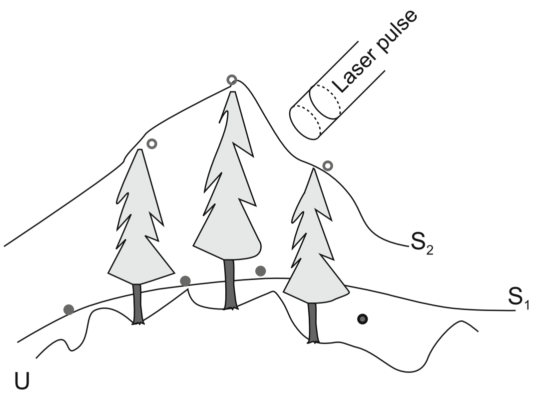


Fig. 9.30 The generation of a DSM/DEM from an airborne laser scanning system generates classification problems. The points belonging to the ground level U must be correctly identified, otherwise they can generate different surface model *(S1, S2)*

Generation: Structural Model Definition

The data for the generation of a DEM are acquired according to different spatial distributions depending on the model that is used and of the DEM final destination.

Surveyed points' spatial distribution can be chosen according to the following criteria:

- regular grids, as typical of photogrammetry;
- irregular grids;
- random and irregularly, as typical of the topographic method;
- ground discontinuity lines, such as valleys, ridges, rivers, lake sides;
- contour lines.

In the most cases, the data participating in the generation of a DEM are not acquired in a single mode, but different distribution criteria can be taken into consideration (Box 9.3).

Box 9.3 Strength and weakness of the Airborne Laser Scanning system compared with aerial photogrammetry and remote sensing in 3D representation

Mapping of surfaces with complex texture

Capacity of volumetric measurement also in dense vegetation. The laser signal penetrates into the vegetation and pulse corresponding to the ground level is recorded

Possibility to differentiate long and narrow details as roads, electric line, costal line, river, etc.

Dense and accurate mapping of steep slope as precipice or building in urban area

Localization and orienteering definition isolated objects as telecommunication antennas

Quick availability of the collected data and fast 3D model transformation

Data accuracy, classification and points identification are factors continuously in progress and often in conflict with the traditional photogrammetry and radar interferometric techniques.

As far as the structural typologies, or final data format, are concerned, in absolute terms, no one is better than the others. Thus, it is necessary to relate the most appropriate data structure to the research and/or application purposes, considering the advantages and disadvantages of each one (Plate 9.5).

The main formats in which a DEM can be supplied are

- *linear or vector format*: the most diffused method for the territory morphology description. It is organized by arcs, or contour lines, connecting points having the same altitude. Contour lines can be easily overlaid onto all the thematic layers of a GIS; they are mainly used as intermediate product, being the input data for the

interpolation algorithms, which lead to a continuous description of the Earth's surface.

- *grid or raster format*: the most commonly used structural model, thanks to its simplicity; the Earth's surface is represented by a regular grid where cells (generally squared ones) are associated to a value representing the height of that portion of territory. All raster GIS use this format for continuous representation of the landscape and its attributes (Fig. 9.31). The limit of the raster format is the persistent discretization of the representation that in most of the cases hides the little discontinuities: no progressive resolution is possible for this type of data; hence if a better description is required, a new DEM having a better resolution must be computed (Fig. 9.32). A quadtree organization is nevertheless possible to transform the grid DEM into a progressive piecewise representation. Advantages of the grid format are the data processing simplicity and the possibility to easily perform multi-temporal comparison and overlapping. A limit is the redundancy of information especially in those areas where little height variations are present: in this case, it has to be considered the eventuality of simplifying the grid size to separate the most homogeneous areas from the most complex ones.
- *polygon grids*: to realize a 3D model of a scene, 3D flat joined polygons (3D geometrical primitives) are often used giving a piecewise representation; a group of primitives, if correctly combined, can describe any shape. A little group of primitives can define simple shapes, as a polyhedron, while more complex objects require a more articulated fractioning of them. Primitives participating in the model are generally simple 3D polygons that can be described in a parametric form inside a grid of polygons. An irregular network of triangles represents a $2\frac{1}{2}$D polygon grid. When the height value, Z, has to be calculated at the generic position (X, Y), the following equation is applied to the triangle that includes the point of the plane:

$$a_1X + a_2Y + a_3 = 0 \tag{9.3}$$

In this case, abrupt height variations create some problems in the representation, due to the bad approximation of the surface granted by a plane. The only solution to improve the local quality of the representation is to increase around that position the number of measured points. Thus, the model is divided into many further portions transforming into a structure where abrupt altitude variations can be better represented.

Geometrical information needed to build the grid of polygons is stored in tables containing the declaration of the coordinates of the vertices and the correspondent connections, whose nature depends on the geometrical relationship that links them. Possible relationships (primitive surface model) can be bi-variate polynomial equations such as linear, bi-quadratic, bi-cubic, bi-quartic equations. The *B-splines* are more complex models used for their geometric continuity characteristics, trigonometric functions, etc.

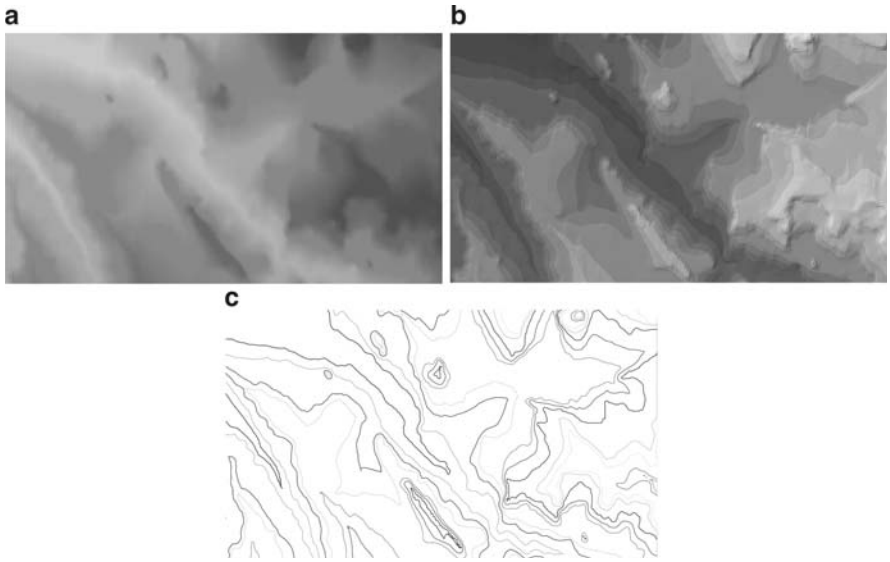


Fig. 9.31 (**a**) Original raster DEM; (**b**) TIN derived from the original DEM and (**c**) contour lines derived from TIN. The contour lines overlapped to TIN are represented in Plate 9.8

- *Triangulated Irregular Network, TIN*: the TIN model represents the Earth's surface by several triangular facets having their vertices coincident with the measured points. As the triangles may have variable size and shape, this structure allows a faithful representation of the Earth's surface, even when strong and abrupt discontinuities are present. Flat zones, in fact, can be represented from a limited number of points and described through big-sized triangular facets; as already recalled, zones where the terrain is more irregular require that triangle

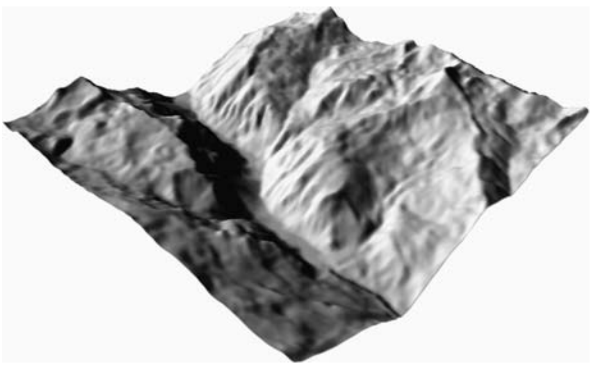


Fig. 9.32 3D view of a DEM in a grid structure in grey scale

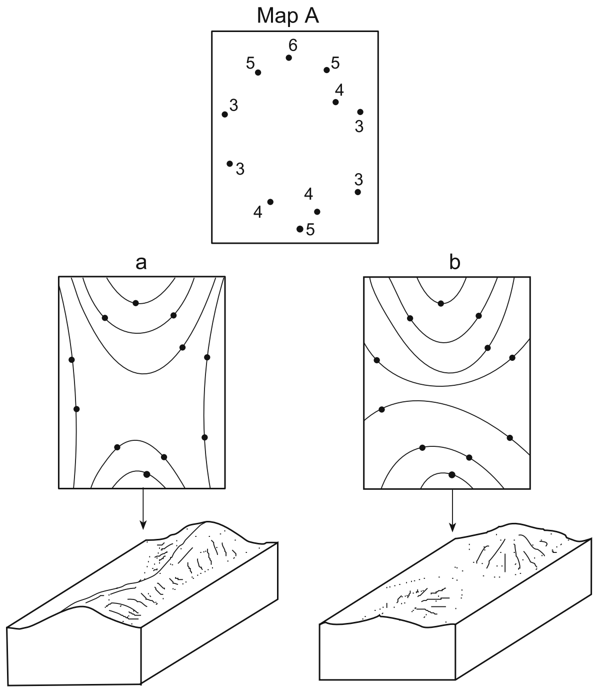


Fig. 9.33 Comparison between two interpolations, starting from the same set of data: the altitude information of the available scattered points in map A can produce very different results (**a**, **b**)

size is progressively decreased by increasing the number of measured vertices. This is the structure mainly used to manage a DEM in a vector GIS. The main limitation of TIN is its difficulty to be managed, having topological relations that must be explicitly declared, that are calculated and recorded on purpose; moreover triangulation algorithms are very complex.

Generation: Interpolation Algorithms

Considering all of the possible data sources, just the utilization of satellite or aerial digital images can directly lead to the generation of a DEM as continuous surface in a discretized form. The most common situation is that requiring point interpolation, aimed at making continuous and opportunely sized, an irregular discrete distribution of measured points or contour lines. The expected results are a grid structure or a TIN. Again, no interpolation algorithm can be defined as best: the choice is related to the DEM structure that the interpolation software can manage, to the ground morphological characteristics and, above all, to the application and required accuracy of the final DEM (Fig. 9.33).

The interpolation process used to generate a *grid or raster structure* computes the unknown height value of the interpolated point as a weighted sum of the height

values of its neighbours (measured points). Weights assigned to each known value are determined as inversely proportional to the distance between the new position and the one of each surrounding measured point. The points selected for the computation can be the nearest *n* (quadrant or octant research criterion), or all the points within a given radius circle (research criterion imposing a maximum radius). To build the grid, i.e. to be able to associate an elevation value to each regularly spaced elementary cell, it is assumed that the altitudes distribution is regular and that discontinuities like ridges, depressions, precipices are absent.

To generate a *Triangulation Irregular Network* (TIN), the most largely used interpolation algorithm is the *Delaunay triangulation*, which provides a univocal association between the measured points and a set of polygons, called *Thiessen polygons*; finally, it generates triangles, different in shape and size representing the described surface by joining the internal points of adjacent polygons. In the new structure, each point becomes a vertex of a triangle, satisfying the Delaunay criterion stating that no vertex is present within the circle circumscribed to a triangle.

Some interpolation algorithms allow generation of a TIN, also using contour lines and raster format DEMs as input data. The latter are generated from measured points, or contour lines, by interpolation operating with numerical methods or approximation methods as weighted distance, Minimum Square, polynomial interpolation, Fourier series.

Elaboration: Modification and Improvement

DEM editing tools are crucial to guarantee good and reliable services inside a GIS. These tools are interactive and allow modifying and refining of an already existing DEM to improve its quality. They permit to search inside the DEM, delete gross errors, add new values, move features, edit the height values, filtering the data.

Digital filters are used to enhance, to smooth the DEM or to reduce the data amount. Enhancement and smoothing filters used during the editing of a DEM are the same used in digital image processing.

Filters are mostly applied to grid structure DEMs; the effects of a *low-pass* filter is to remove the details and hence make the DEM more homogeneous; while a *high-pass* filter has the opposite effect of enhancing surface discontinuity.

Sometimes the filters are used to eliminate redundant data, saving space in the mass memory of the PC and reducing the processing time. They can also be used to reduce the DEM resolution, especially before converting one structure into another. Two adjacent DEMs can be mosaiced, while two or more DEMs partially overlapped can be merged together to form a unique model.

Another possible operation and processing is the conversion of an existing DEM having a certain structure (i.e. grid) into another structure (i.e. TIN). Grid models are easier to be edited, as operations refer to a single cell; TIN editing, instead, requires tools that are more powerful because each intervention changes its topology (Box 9.4).

Box 9.4 Terrain morphology modelling by DEM interpolation and some interpolation methods

Terrain morphology reconstruction by DEM interpolation	
$Z = f(X,Y)$	where: Z: altimetric coordinate X, Y: planimetric coordinate
DEM interpolation methods	
Pondered mean interpolation	It mediates the Z values of the points around the considered P point, giving to each of them a higher weight in proximity of the topographic position X, Y of the P point.
Linear superficial interpolation	$Z = a_1X + a_2Y + a_3$
Bi-linear interpolation	$Z = a_1XY + a_2X + a_3\,Y + a_4$
Reticular interpolation	$Z = a_1X^2Y + a_2{\cdot}XY^2 + a_3X^2 + a_3{\cdot}Y^2 + + a_5{\cdot}XY + a_6{\cdot}X + a_7{\cdot}Y + a_8$

DEM Interpretation: Analysis and Information Extraction

GISs are usually equipped with appropriate tools aimed at processing DEM to extract further information about the geometry of the territory. There are two possible ways to follow:

- simple visual interpretation of the DEM;
- quantitative analysis.

The result of these approaches is information useful for environmental impact analysis, for the definition of models concerning the potential erosion or to simulate the water flows (Plates 9.6 and 9.7).

The aim of DEM analysis is to derive morphometric parameters that can be classified as:

- *general*: this is the case with:
  - slope maps, raster data where the cell value represents the angle between the horizontal and the local gradient of the DEM (referred to the local tangent plane);
  - aspect maps, where the cell value represents the angle (azimuth, clockwise) between the north direction and the local gradient (planar components). Usually nine classes are considered: one referred to the null value (flat areas), the other height to the main direction sectors (north, northeast, east, south-east, south, south-west, west, north-west). These maps are used in the geological and environmental fields;

  - perspective views, viewshed analysis, height profiles and volume calculations can be also be carried out through this family of tools.

- *specific*: morphometric parameters directly related to the surface hydrological characteristics. Among the point characteristics can be highlighted: depressions, peaks and passes; among the linear ones: drainage channels and watershed boundaries; among the polygonal ones: drainage basins subtended by a closing section. These parameters can be investigated and used as input layers in hydrological modelling of superficial flows simulation, or in flood forecasting or in geomorphological modelling. A DEM can also be used to calculate terrain volumes exceeding a reference surface (e.g. a cutting plane). This is obtained as the sum of the volumes of the several prisms built on a base that can be the grid cell or the triangle planar projection.

Visualization
DEM visualization can be

- *static:*mostly used to communicate modelling results by printed copies or digital support;
- *interactive*: enabling the user to explore the product and possibly modify and refine it.

DEMs visualization conventional modes are the contour lines and the hillshade (flat representations) and the virtual 3D representation.

Applications

Application fields of DEMs are several: topographic survey, photogrammetry, civil engineering, planning and resource management, Earth and environmental science, military applications, remote sensing and going into details:

- *civil engineering*: landscape planning and visibility analysis, planning of strong environmental impact infrastructures such as roads, dams, bridges;
- *geology*: slope and aspect maps for geomorphologic studies and assessment of landslide and erosion phenomena;
- *agrometeorology*: slope, sun-exposure and aspect maps for the study of cultivated plants phenology and production;
- *hydrogeology*: hydrographical network extraction and hydrological basins geomorphological characterization;
- *remote sensing and cartography*: radiometric data, orthoimages, 3D images, thematic maps interpretation.

9.9 GIS Evolution

GIS concept is in continuous evolution thanks to the increasing availability of new analytical functions that offer user-friendly interfaces to operate more and more efficiently and economically.

Until now, the attention has been paid to some fields where GIS is used for analysis and description of the territory. The next expected step is to the possibility of using GIS to simulate human intervention, or disaster events to forecast their impact, in order to answer to questions about what, where and when. Time must be considered as the fourth dimension of geographical data. For example, in case of leakage of pollutants from a chemical plant, *the riskiest wells, irrigation channels, rivers*could be identified *and the pollutants permanence determined.*

9.9.1 GIS Object Oriented

Object-oriented models, suitable for spatial management, have been developed since the first half of 1980s.

The most commonly used database model for GIS is the relational one; due to its simplicity and thanks to its standard language, SQL (*Structured Query Language*), it is considered more efficient than the hierarchical and reticular (network) ones. Hence, GIS non-spatial data organization mainly refers to relational structures, while geometric data are managed in topological structures.

The relational model presents some limits:

- it is inefficient when the database contains a huge amount of information;
- it is quite inadequate to process complex objects like the GIS spatial objects;
- it lacks the appropriate mechanisms for data structuring: e.g. thematic and geometric data, referring to the same object, are organized in two different management systems.

With the development of object-oriented programming languages, also called fourth-generation languages, the trend is the unification of the geometric and attribute structures. The new approach, called o*bject-oriented data modelling*, is suitable for spatial applications to describe the elements in the real world and the relative properties as objects, in order to support the processing of complex geometric entities.

The *field-based approach* models the real world as a *non-empty* space composed of area units, each described by its own attribute data.

The *object-oriented approach* conceives the real world as an *empty* space filled by single ground objects. In this case a process of abstraction starts, concerning the editing of the ground objects (*what*), their localization (*where*), the moment (*when*) and the relations among them. In order to simplify the abstraction process, one of the three elements is kept fixed, one is pre-defined and the third one is measured or observed.

Before organizing spatial data in a system, these three elements have to be identified and formalized. The spatial database represents the objects of the real world as they appear in a specific application. Semantic aspects of each ground object can be analysed geometrically and thematically; the first includes topology, shape, size and position, while the second one refers to the non-spatial attributes of the object.

From this point of view, the term *object* is used with a wider meaning than the elements reported on a topographic or thematic map. In fact it describes an entity in a single moment, specifying its individuality and its behaviour (rivers, fields, soil units, etc.).

The main terms and concepts used in the object-oriented approach can be grouped in a *model structure* and a *development structure*. The first one includes the object identity and the four abstraction mechanisms that characterize it: classification, generalization/specialization, aggregation and association.

9.9.2 Decision Support Systems (DSS)

GIS-advanced utilization is increasingly evolving into Decision Support Systems. In this case GIS is no longer used just for analysing data and simulate events, but it is exploited to choose, within a group of possible alternatives, the optimal solution for a certain problem. Some functions of this kind are already present in different GIS (i.e. the calculation of the optimal path between two points, the choice of an area using different query functions, etc.), and in some cases they are developed to satisfy specific needs of the user. Decision Support Systems assist the operations flow that from the real-world observation leads to the planning of appropriate intervention onto the territory.

Decision Support Systems (DSS) host very sophisticated (and expensive) processing tools, and they are conceived to answer the question:

what is going to happen if...?

They can simulate the world, create potential scenarios and suggest a solution to the decision maker. Some operative experiences have been made in the field of the *Environmental Impact Assessment* (EIA), where appropriate DSS are structured to analyse and process environmental data to investigate the effects of human activities and buildings. Analysis is transparent and all the processing steps can be documented and stored, allowing their easy recalling aimed at fast repetition of defined workflows or at their improvement.

Decision Support Systems can be thought of as a coordination tool between different professionals and competences and a communication instrument between the technical and the social world, making decision processes transparent, flexible and interactive.

These systems are required to be able to run algorithms that model the examined phenomenon and to represent the calculations results continuously giving the operators the possibility of following their evolutions or simulating different scenarios. Modelled phenomena are considered to develop over terrestrial backgrounds that can be changed according to the investigated different situations. In some cases,

also sophisticated 3D managements are required to enhance those aspects of the phenomenon that are difficult to observe with a bi-dimensional visualization.

Practical examples are emergencies simulations and management, such as hydrogeological risk, fires, environmental pollution, aerial traffic control, tactical situation management, selection of the most suitable area to position a ground survey network. Control Room management is one of the most suitable contexts where the DSS can operate. In fact, monitoring of a technological network referring to production processes, mobile phones management, radio-bridges, electric/railway/road transmission, etc., for civil security and protection, is an appropriate benchmark to test this technology.

Decision Support Systems include the following elements:

- software library;
- archive structure;
- data input and output;
- data management in a coherent context;
- management of the operative dynamical structures overlaid onto the basic theme layers;
- management of the spatial functions representing the effect of the evolution of specific phenomena models;
- extraction of the information to be processed;
- distributed structure of these systems and related complex cooperative and network functions.

GIS commercial software generally satisfies just a part of the above-mentioned features, but they often are not flexible enough to be directly adapted to the specific needs of a decision support system. Nevertheless, there are some new products which may overcome such limitation.

9.9.3 Expert Systems (ES)

Another trend which might improve GIS spread, even and especially at a local level (little professional realities), is the *Virtual GIS* (VGIS), mainly characterized by the availability of simplified user interfaces also allowing non-qualified personnel to easily manage analysis functions in a complex GIS. The goal is to equip small technical offices with a tool not only devoted to the geographical databases consultation but also to a simplified processing (basic operations) of the data. More complex and less frequent operations remain accessible only by specialized users.

This instrument should allow the user to forget about GIS design problems and to concentrate on specific tasks and practical needs regarding a particular purpose or product.

Geometrical models defined in most GIS and spatial concepts resident in the users' mind are usually quite different. GIS can differ with respect to the adopted geometrical model and, consequently to the spatial concepts; furthermore, each problem is faced exploiting only appropriate operations related to the chosen model.

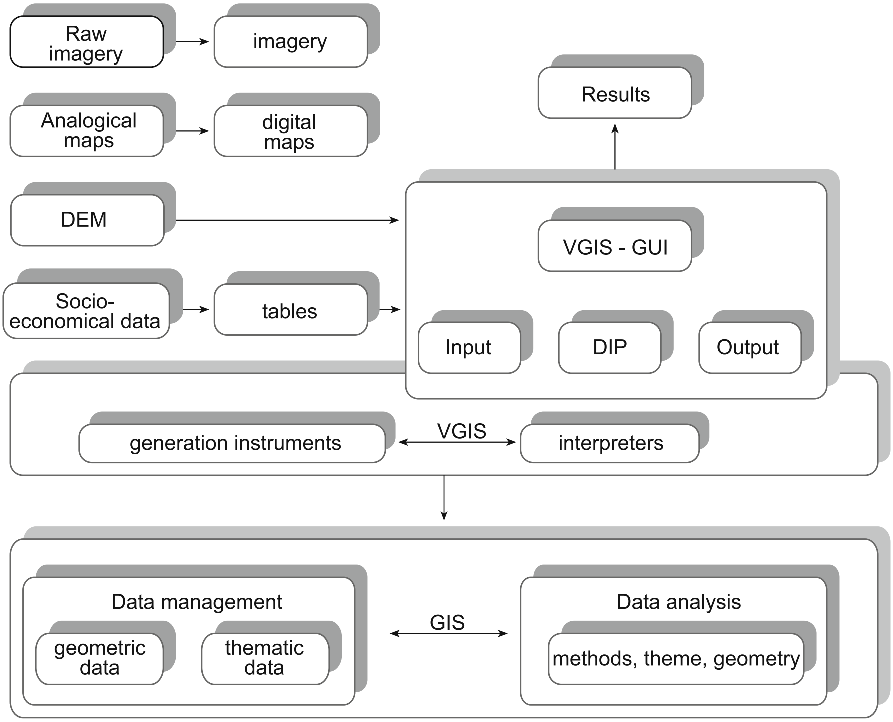


Fig. 9.34 The virtual GIS as proposed by Ehlers, 1995. GUI: Graphical User Interface, DIP: Digital Image Processing,. DEM: Digital Elevation Model

VGIS are a trial of integration of different spatial concepts helping to make the GIS nearer to the user.

Virtual GIS are conceived as a possible approach in automatic management of huge amounts of data inside a GIS. They use Expert Systems that are complex algorithms able to acquire and formalize knowledge to perform self-managed operations (e.g. automatic interpretation of images acquired by different platforms).

Virtual GIS processors (inferential) design and develop workflows based on knowledge, and they offer a user-friendly *Graphical Users Interface* (GUI) oriented to applications. This approach leaves the user free to concentrate on the objects rather than on the data and their management (Fig. 9.34).

The key, making virtual GIS independent from the traditional GIS, is their specific orientation to user operability, by facilitating data analysis and information extraction, and providing results in a synthetic and exhaustive format. Each task of the process can be divided into elementary blocks acting as the atoms of a molecule. The first step is the development of opportune interfaces permitting data (mainly images) processing even to those specialists not aware about of GIS internal organization.

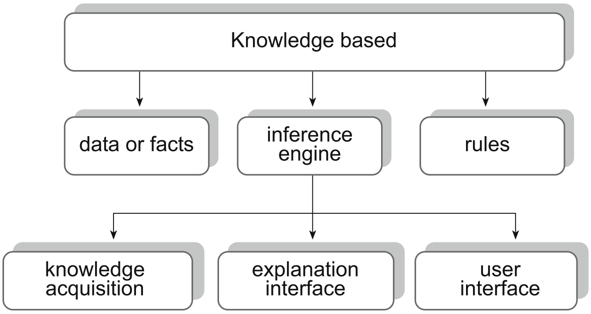


Fig. 9.35 Expert system components

Thus the virtual GIS aims to reach a complete modelling to obtain an automatic data interpretation in a GIS context. Its widest goal includes not only the modelling of all of the data (attributes, maps and images) and of the interpretation processes, but also of the applications; this grants that the user can operate and decide in a more and more objective way. In order to reach this declared goal, adopted algorithms have necessarily to refer to the fields of Artificial Intelligence or of Expert Systems, better known as *Knowledge-Based Systems*.

An Expert System organizes the knowledge on three levels including data (or facts), the rules for their management and the control procedures generated by the inference processor (or *shell*). The shell is the place where the choices and actions based on the suggestions coming from the experts are generated. The inference engine is a system using the knowledge to get to some conclusions (Fig. 9.35).

The user develops his own conceptual model concerning the expected result exploiting a system operating according to many logical sequences based on his experience. In current GIS this ideal division is up to the operator who, hence, has to possess a deep knowledge about all the system functions he is going to use; in virtual GIS this step is completely automated thanks to the Expert Systems (Fig. 9.36).

Virtual GIS basic components are

- a visual language interface;
- an interpreter of the process workflow (59.36);
- a GIS, conceived as a black box, where interfaces lead the necessary operations;
- a process workflow designer/builder.

The VGIS interpreter converts and splits the user's workflow into two steps:

- generation of the required GIS elementary, or basic, functions;
- mapping of the results in a traditional GIS way.

This scheme shows the will of creating specific workflows by assembling basic functions.

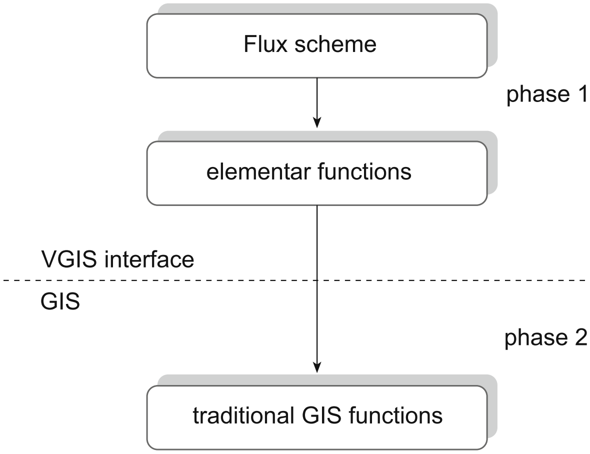


Fig. 9.36 Process workflow of a virtual GIS

The workflow builder aims at facilitating user's operations by developing secure and user-friendly interfaces. The user has just to define the product and the data source waiting for the automatic system response mainly represented by a personalized flowchart equipped with exhaustive explanations. The user can modify, at any moment, the proposed solution. Process workflow designing and building is organized in blocks representing the elementary functions of the GIS concerning information generation exploiting available meta-information, description and data about the study area specified by the user.

The process workflow is based on backward chaining Expert Systems where rules refer to two components:

- *objects models* (metadata), describing and identifying the data in a detailed way;
- *process models*, describing the VGIS functions application.

9.9.4 Role of the ES in Image Interpretation and Classification

As application complexity grows, image interpretation and classification show some evident limits, which can partially be overcome by introducing Artificial Intelligence (AI) techniques. Complexity of applications depends on information heterogeneity, number of decision classes and required performance level.

An effective solution must be based on tools able to *simulate*knowledge processes generated by the experts according to their skill to manage real-world complexity by *abstraction*, *generalization* and *approximation* interdependent processes.

Currently it is possible to state that a knowledge-based system, or *Expert System*, if properly *trained*, can perform complex interpretation and classification functions referring to the following elements:

- the choice of an appropriate *language* to represent the knowledge;
- the development of an *inferential engine* able to generate the results related to the data and the represented concepts;
- an appropriate *strategy* for knowledge acquisition.

The solutions depend on the specific application domain and type of problem to be solved. What distinguishes a system based on knowledge with respect to any computer program is, most of all, its architecture that includes an explicit codification of facts and concepts relative to the considered problem.

The Expert Systems perform a *transparent* approach to the problem; they have, in fact, the property to explicitly document the required knowledge, the concepts and the casual relationships among them in order that the result can be qualified with respect to the *how* and the *why* it has been obtained.

9.9.4.1 Knowledge Representation: Role of the Fuzzy Logic

During the 1980s, there were a strengthening and an increasing spread of knowledge representation methods for Expert Systems, essentially based on formalisms derived from the first-order Classic Logic and on schemes derived from it such as semantic networks, frames, Horn's clauses (Cercone & McCalla, 1983).

First-order logic is a formal representation and symbolic processing system that already contains the tools to describe knowledge portions or specific concept by means of symbols referring to precise syntactic rules. An inferential scheme builds symbolic processing sequences for the inference of new facts.

Expert Systems in the past have been based on symbolic manipulation languages like LISP, Prolog or domains (*shells*) with representation and inference pre-defined schemes.

Conventional logic formalism has Boolean and normative character. According to a Boolean approach, a language sentence, expressing, for example, object belongings to a given class, can assume just true or false values; logical connections for predicates and inferential scheme aggregation have a pre-defined meaning independent of the context and experts' subjectivity.

The knowledge of a problem is often incomplete or not precise; the experts use natural language sentences to communicate intrinsically vague and undefined concepts able to discriminate a true or false value. Knowledge representation on the contrary must be based on formalisms able to *imitate* the experts' knowledge processes in formulating qualitative descriptions useful to resume the real-world complex details necessary to take rational decisions.

In literature are suggested many solutions to the problem of knowledge representation, mainly based on the alternative use of non-standard logics representing qualitative, intrinsically uncertain knowledge and reasoning approximated schemes. Among the most widespread models are Mycin's Model, Dempster Shafer Theory, Rough Sets, Fuzzy Logic and Bayesian Models (Smets et al. 1988).

Fuzzy Logic is based on possibility theory (Zadeh, 1973) and derived from the *Fuzzy Sets* theory (Zadeh, 1965); it is a powerful and flexible formalism to face heterogeneous data modelling and classification problems and to imitate the experts'

ability in taking deterministic rational decisions, in high conceptual complexity contexts (Zadeh, 1981).

A representation scheme based on fuzzy logic uses predicates having continuous true values in the interval [0, 1] and specializes inferential scheme semantics in function of the context and subjectivity of the field experts.

In a typical interpretation/classification fuzzy scheme, it is supposed that for each *observable element* (for example an image cell or pixel), it is possible to distinguish different values related to its probability to belong to the considered classes. These values derive from an abstraction process realized by the experts when describing their assessments about the objects.

For example, let us consider the observable quantity *ground elevation* with a (0, 200) interval expressed in metres. The experts can describe the interpretation of the *elevation* data in terms of *high*, *medium*, *low*, operating a sort of abstraction on the data numerical multiplicity. Although these classes are conventional and do not correspond to terms really used by the experts, they express the right information granularity on which the data interpreter has to base its decisions.

It is assumed that the abstraction process is possible in nature, and these classes can be represented by fuzzy sets defined in the numerical interval of the corresponding observable quantity.

Each linguistic description like *high ground elevation*, which will be the predicate in the interpretation scheme, is associated with a probability distribution. A precise altitude value will correspond to a value of probability to interpret the *high* altitude. Each *observable quantity*'s descriptions is aggregated by logical operators (i.e. AND, OR) whose semantics can be specialized according to context considerations.

Fuzzy logic formal structures hence allow getting the knowledge intrinsic vagueness, typical of many application fields.

Knowledge representation formalisms, like the ones based on fuzzy logic, can hence fill the gaps between mind representation of the concepts and mathematic models, contributing to the definition of reliable, transparent and accurate intelligent systems.

Anyway some experiences, mostly related to geographic and environmental data analysis, highlighted some limitations in the use of only one knowledge representation formalism. These basically depend on:

- the number of observable quantities, to be evaluated in a combined way to take intermediate or final decisions (i.e. more spectral bands of an image);
- the necessity to discriminate among several undefined cases (i.e. mixed pixels) due to the presence of a high number of decision classes.

No model (Mycin's Model, Dempster Shafer Theory, Rough Sets, Fuzzy Logic and Bayesian Model) is universally valid; thus, no model can satisfy at the best different requirements of complex domains. According to this concept, the current trend is to consider as complementary these representation models, previously considered as alternatives, and to focus the interest on hybrid architecture systems

where two or more knowledge representation models cooperate for the same (or different) goals.

Hybrid intelligence systems thus are based on the integration of two different representation models, for example symbolic and connective (neural) ones (Rumelhart et al., 1986), or, in the symbolic domain, on the integration of different uncertainty processing models (i.e. based on probabilistic theories). The different models integrated in hybrid system architecture cooperate at different levels, preserving mutual advantages and overcoming some limits that appear if used one at a time (Binaghi et al. 1993).

9.9.4.2 Knowledge Acquisition

The knowledge acquisition step is the limiting factor in the definition of an Expert System. Currently there are different approaches to the problem:

- psychological and social science concepts and method integration (Neale, 1988);
- automatic generation (Machine Learning), which has defined strengthened several methods in many applicative domains (Michalsky, 1994).

Among the most diffused psychological methods dealing with past event analysis, besides the introspective ones, it is worth referring to the *structured interview technique* adopted during the first steps of interaction with the experts. It is devoted to verifying the project feasibility and to acquiring the technical terms and defining the basic problems.

The most common interview technique can be extended to all of the steps of knowledge that develop through a structured sequence of questions to ask to the experts. The sequence and the type of the questions are led by the nature of the problem and by the formal model this knowledge has to be codified in. Direct experience in the use of structured interview techniques demonstrated that they integrate well with the choice to represent knowledge by fuzzy logic.

A diffused automatic learning method, especially suitable for the acquisition of knowledge by Expert Systems applied to classification problems, is *empirical learning*or *induction by examples*. This approach requires

- a set of supervised samples referring to specific situations about which the experts have already expressed an opinion, or selected a decision;
- an induction mechanism capable to abstract from specific cases and generate rules on which the decision making is based.

Knowledge acquisition and representation strategies are strictly related and one conditions the other, affecting the efficiency of the system. Considering both these issues suggests that Artificial Intelligence can offer a valid tool for supporting complex decisional activities and it plays a precise role, sometimes exclusive, in facing problems of heterogeneous data automatic classification in complex real fields.

Strong and efficient solutions are found by integrating these several approaches. Since, as already noted, there is not a universally valid method, all the methods add their own specific value; the synergic cooperation among the different strategies interacting inside hybrid architecture systems is the basic chance offered by Artificial Intelligence techniques to solve operational problems.

9.10 Error, Accuracy, Precision and Tolerance

The problems related to data errors, like the lack of metric or semantic accuracy, affecting Spatial Information Systems have been widely underestimated for years; experiences suggest paying great attention to these issues, in order to improve the GIS reliability.

GIS has the capability of linking and integrating different data based on their location independently from their nature or source. Nevertheless, the importation of new data into the system often suffers from possible, not necessarily evident, errors that sometimes can even be forecast.

The crucial point is that possible errors must be known and their occurrence possibly foreseen in order to reduce as much as possible their effect on the result. The need of an accurate knowledge about errors leads to a deep objective evaluation of the GIS potential limitations that must be pointed out to avoid misleading interpretations of the obtained results.

9.10.1 Definitions

In order to better understand the problem, it is fundamental to introduce some definitions taken from the errors statistic theory:

- *Accuracy*: it is the degree of similarity between the measurement and the real value of the observed quantity; it can be improved by an instrument or procedure calibration. It is calculated as the difference between the mean value of the measurement and its expected theoretical value; often it is indicated as bias. In the case of a GIS, it can be defined as the degree of similarity between measurements (or information) derived from GIS data and the surface reference (or values widely and previously accepted). While dealing with metric measurements, it is always possible to separately consider horizontal and vertical accuracy. The level of accuracy required for specific applications can vary widely; as the generation of accurate data is a very expensive and difficult operation, the final goal and the required accuracy must be seriously considered. Statistically speaking, the bias (in altimetry as well as in planimetry) tends to zero as the accuracy increases. The geometric accuracy depends on the map scale.
- *Precision*: the degree of reciprocal consistency among the values of repeated measurements of the same quantity, which have to be performed respecting all of the following conditions:

– same measurement method;
– same measurement instrument;
– same observer;
– same place;
– same working conditions;
– repetition within a short time period.

Precision defines an intrinsic quality of the adopted instrument, and it defines the degree of dispersion of measurements around their mean value. In a GIS, the level of precision can vary widely in function of the application it refers to: topographic measures of a building require higher data precision than the definition of urban area perimeter. High precision does not necessarily mean high accuracy and vice versa; but they have in common the high costs to be obtained. In probability and statistics, precision defines the measurement repeatability, and it is higher as the square root of the variance (MSD), called Standard Deviation, gets smaller (see Box 8.5).

Box 9.5 Types of error that can occur creating a geographical information system

Type of error			
Evident error sources	**Natural variation or origin measurement**	**Elaboration**	**Errors propagation**
Data updating	Positioning accuracy	Numerical	Propagation
Remote sensing data availability	Content accuracy	Topological analysis	Cascade
Map scale	Classification generalization	Digitalization	
Observation frequency	Source of data variation	Geocoding	
Deducing (extrapolation)			
Format			
Accessibility			
Cost			

- *Data quality*: defined by the combined effects of accuracy and the precision of the data. Data quality requires to be stated by declaration of the accuracy and precision data and also of the processing steps followed for their determination.
- *Error*: anything that negatively affects accuracy and precision defines the data error. It can be gross, systematic or casual.

- *Tolerance*: it has to be considered the real reference value to define the map data quality. It is usually set equal to two × MSD. About 96% of the measurements are included in the range mean ± 2 × MSD (admitting a normal distribution of the observed variable). For example, the tolerance of 1:100,000 cartography is 40 m.

9.10.2 Types of Error

Errors can have different origins, but positioning error is the most evident and also the more usual. Nevertheless there are many other errors affecting the attributes; they can be conceptual and logical, and they can cause serious damage to the system, although they are not so evident:

- *position accuracy and precision*: refers to both planimetric and altimetric elements. The scale plays a determining role conditioning the applications of the map. A point on a map does not have a unique fixed position, but a *probable* position within a certain area. Accuracy and precision values are strictly related to the nominal map scale, and the possibility to zoom an image in a GIS system does not modify this basic concept;
- *attributes accuracy and precision*: non-spatial data often suffer from inaccuracy and imprecision due to an incomplete description of the represented phenomena or elements;
- *conceptual accuracy and precision*: let us consider the case in which a GIS contains a classification derived from remotely sensed data. If inappropriate land cover classes have been defined for a certain area, a conceptual error is committed. Admitting the selected classes are appropriate, however, commission or omission error can occur. The false indication of a tributary river position is a typical commission error; not recognizing it, instead, is an omission error;
- *logic accuracy and precision*: the information collected in a database can be organized according to different logics that sometimes can be incorrect or not appropriate. A GIS usually cannot alarm the users about an incorrect use of the data. An Expert System that processes data inside a workflow, performing queries and obtaining answers in a complex structure and logic, is sensitive to this kind of error.

9.10.3 Sources of Error

The possible sources of errors that may affect a data set quality in a GIS can be many and of different nature. In Box 9.5, a list of the different types of errors is reported.

9.10.3.1 Evident Sources of Error

Some of these errors can be avoided because they are related to operator's superficiality, weak attention or just to the filling of missing data with substituting ones

having a lower or different quality. Other errors depend on some choices external to the GIS manager who is responsible for assessing the costs/benefits of accepting or rejecting some kind of errors:

- *data updating*: structured data within the system can result in inappropriateness for current applications as generated in the past when both object and methods conditions were different. This difference can often generate inconsistencies. For example a map that is not up-to-date is not suitable to represent recent effects caused by erosion, deposits or other geomorphological processes;
- *remotely sensed data availability*: acquisition modalities of satellite or aerial remotely sensed data often produce incomplete data. Problems can be a lack of data inside an image (cloud cover, acquisition errors), or a lack of images in a time series (image unavailability in the requested dates). Both these situations cause a non-homogeneous availability of imagery;
- *map scale*: the possibility to represent details in a map depends on its nominal scale; the scale affects data type, quantity and quality. Maps at a lower scale too often are used to deduce data and information at a higher scale;
- *data observation frequency:* a small number of observations in a certain area can lead to performing inappropriate spatial analysis. For example, if the contour line interval is 40 m, it is not possible to declare an accuracy of results higher than this threshold. Considering that contour lines are derived from a dispersion of point data, it can be stated that the denser the point distribution, the higher is the accuracy of the resulting contouring;
- *data deduction*: when data are not directly available, they can be indirectly assessed and/or represented (extrapolation). Derived (secondary) data must be generated paying a lot of attention and trying to qualify their accuracy, which is to model error propagation. Data generated by deduction (spatialization) are certainly needed for many applications, but they do not ensure an accuracy level equal to the original ones they are derived from;
- *data formatting*: digital information formatting is required for their efficient archiving, transmission and processing. Some errors can be introduced even during this operation. Original data alteration is always possible while converting between reference systems, or during data format conversion (raster to vector or vice versa), during image radiometric resampling. Continuous conversions from one format to another can produce an effect similar to the repeated photocopying of documents, i.e. error propagation and hence data degeneration;
- *data accessibility*: often data and the information exist, but they are not directly accessible. Restrictions due to security, laws, competition between institutions and economical factors limit the availability or the accuracy of the data, which can be provided incomplete or with a degraded resolution;
- *data cost*: obtaining and converting data can be a very expensive step; maintaining and managing a ground survey network can be a very expensive choice. Costs must be considered as the result of a compromise among the collected data types, their quantity and accuracy. A GIS manager must carefully consider this.

9.10.3.2 Errors Depending on Natural Variation and on Measurement Methods

These errors concern the object position. They can be related to the map quality, to the effects of instrument mis-calibration or to the quality of a land use/land cover classification:

- *map geometry related errors*: this measures the deviation between the true object position and its position on the map. Object position has its own accuracy including errors coming from digitization, scanning and conversion errors;
- *map content related errors*: related to the assignment of polygons to the correct class. If a fir is classified as a pine, it can be irrelevant for some applications but it can be unacceptable in the forestry field;
- *instrument related errors*: instrument calibration and surveying conditions can vary during field campaigns. These variations have always to be considered and evaluated. For instance, the season of acquisition affects the illumination conditions of an image; the accuracy of a differential GPS positioning in the same area depends on drought or wetness of atmosphere.

9.10.3.3 Errors in the Processing

These errors are more difficult to be detected by the user as related to the adopted hardware and software. Often the user does not know the GIS algorithm details; he just runs it to produce a result. This superficial approach can lead to a wrong choice or a wrong use of the algorithm producing some errors:

- *numerical errors*: not all the processors can face the same complex computations, and the possibility of generating certain results is strongly conditioned by this. For example, image processing generates results that can be controlled in mathematical terms only by highly skilled professionals, usually different from geomatics experts. Sometimes the error can even be inherent the basic components of a computer, e.g. the chips;
- *errors in topological analysis*: logic errors can generate incorrect data processing and topological analyses; the data can be non-uniform and suffering from variations; thus, the integration of separately generated thematic layers can lead to some problems related to the correspondence between geometric features;
- *classification and generalization*: the definition of thresholds controlling the assignment of an element to a given class or the visualization of a too small area are typical operations that lead to these errors;
- *errors in digitizing and georeferencing*: while acquiring points, lines and polygons from a map placed on a digitizer and during image georeferencing, there are several opportunities of committing errors (Figs. 9.37 and 9.38).

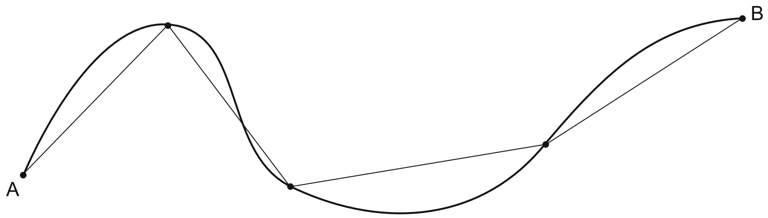
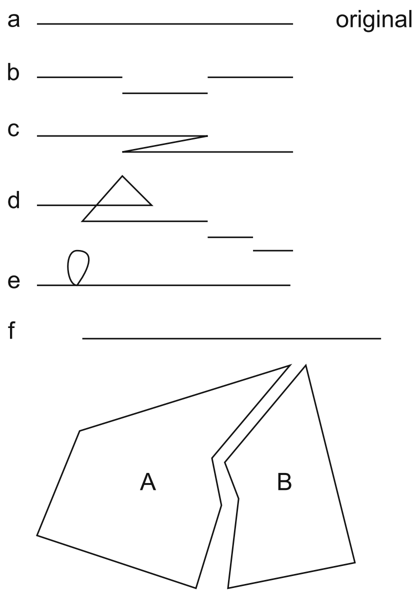


Fig. 9.37 Digitization errors: a curve becomes a polyline

Fig. 9.38 Common digitization errors



9.10.3.4 Error Propagation

A GIS is formed by data sets having different nature and coming from different sources with a very low probability of having the same accuracy and precision. Moreover transferring the data into digital format can introduce further errors. According to this situation, GIS can result globally quite inaccurate and imprecise, as errors inside original data suffer from some transformations along the processing flow. Transformation affecting original data errors can be

- *propagation*: this occurs when incorrect data are used to perform a single operation generating new data; for example when a control point wrongly measured is used to register images or maps. To reduce this event, maps at the highest scale possible should be used to acquire Ground Control Points;
- *cascade*: this occurs when a propagation error is repetitively and in uncontrolled way transferred to many thematic layers. This is the case with an incorrectly

georefenced image used as reference map for other images registration. This type of error can be additive or multiplicative and depends on how the data and the information interact.

Cascade and propagation can affect all the other errors considered previously.

9.11 Metadata and Data Quality

When data accuracy and precision are unknown, it can be very dangerous to use them in a GIS. Apparently carefully collected and organized data can be imported or related in a database in an unsuitable way with respect to the application addressed. Thus, it is important to know detailed information when acquiring data to be loaded in a GIS.

For this reason, metadata are strongly recommended, in order to give an Identity Card to the data, qualifying them.

Metadata certify the global quality of data sets; a data set is defined as a logic group of one or more objects physically, or electronically, stored. The components defining an object are, for example, in the case of a book: title, authors, date and publication place, synthetic describers, abstract, etc. Metadata include elements useful to identify the data, the owners and the access conditions to the data, to declare the reference system, the content and the previous processing, accuracy and precision (Box 9.6). Some elements of the metadata, like accuracy and precision, require that a particular analysis be carried out to obtain quantitative indices that represent them.

In Europe the CENT/TC 287 (*Committee of Normalisation, Technical Committee* 287) *Geographic Information* has been working for years on the documentation Standard; in 1999 the Rule collection about the geographical metadata, indicated as pre-Standard CENT/TC 287, was published.

In 1994 in the United States, the document *Content Standard for Digital Geospatial Metadata*, edited by the *Federal Geographic Data Committee* (FGDC, 8th June 1994), was published. Moreover, organization rules that had been defined and geographical data documented according to FGDC standard were made available.

The CENT/TC 287 document is divided into sections:

- *data identification*, information needed to univocally identify the available geographical data (geodata sets);
- *general description* of available geographical data;
- *quality* of the database;
- *spatial reference system* used to position the geographic objects;
- planimetric and 3D geographical data, and temporal *extensions*;
- data *definition* to facilitate the comparison among geodata sets and to distinguish between different classes of geographical objects, between geographic objects belonging to the same class and to describe the relationships between classes and objects;

Box 9.6 Example of metadata (Standard ISO:19139) referring to Snow Covered Area daily map products created by NASA-MODIS Snow and Sea ice Global Mapping Project and distributed in an INSPIRE compliant Spatial Data Infrastructure (courtesy of AWARE and IDE-Universe Projects, supported by European Commission)

Identification info	
Title	MOD10A1 Snow_Cover_Daily_Tile 2007-281
Date	2007-10-13
Date type	Creation
Edition	Collection 5
Individual name	Dorothy K. Hall
Organization name	NASA/Goddard Space Flight Center
Position name	Responsible official MODIS Snow and Sea ice Global Mapping Project
Voice	301-614-5771
Facsimile	301-614-5644
Delivery point	Cryospheric Sciences Branch, Code 614.1
City	Greenbelt
Administrative area	MD
Postal code	20771
Country	USA
Electronic mail	dorothy.k.hall(AT)nasa.gov
OnLine resource	http://modis-snow-ice.gsfc.nasa.gov/
Role	originator
Presentation form	imageDigital
Collective title	MOD10A1 h18v04
Abstract	MODIS Snow and Sea ice Global Mapping Project provides a variety of snow and ice products from the MODIS sensors, and the products are available at a variety of spatial and temporal resolutions. The daily MOD10A1 and MYD10A1 snow products are tiles of data grid in the sinusoidal projection at 500 m spatial resolution
Purpose	The MOD10A1 product (tile: h18v04) has been produced by the MODIS snow and sea ice Global Mapping Project, and then re-projected to UTM32WGS84 by NASA MODIS Re-projection Tool ver3.3 to provide snow cover area information over the Alps to the authorized users of the AWARE EC-project
Credit	NASA MODIS Snow and Sea ice Global Mapping Project
Point of contact	
Individual name	Dorothy K. Hall
Organisation	NASA/Goddard Space Flight Center
Position name	Responsible official MODIS Snow and Sea ice Global Mapping Project
Voice	301-614-5771
Facsimile	301-614-5644
Delivery point	Cryospheric Sciences Branch, Code 614.1
City	Greenbelt
Administrative area	MD
Postal code	20771
Country	USA
Electronic mail	dorothy.k.hall@nasa.gov
OnLine resource	http://modis-snow-ice.gsfc.nasa.gov/
Role	originator

Descriptive keywords	Remote Sensing, Hydrology (discipline). MODIS Snow and Sea Ice Global Mapping Project , AWARE project , Thematic map , Environment , Climatic change , Water , Snow (theme)
Access constraints	intellectualPropertyRights
Use constraints	intellectualPropertyRights
Other constraints	The AWARE project has been authorized to distribute re-projected MOD10A1 data from the MODIS Snow and Sea ice Global Mapping Project for the use of them within the project framework
Spatial repr. type	Grid
Topic category code	imageryBaseMapsEarthCover

Extent			
Geographic bounding box			
		North bound latitude 50.16	
	West bound longitude −1.7		*East bound longitude* 15.58
		South bound latitude 39.65	
Temporal			
Time Position	2007-10-08		
Distribution info			
Distributor			
Individual name	Monica Pepe		
Organization	CNR-IREA		
Position name	AWARE project Technical Manager		
Voice	+39 0223699550		
Delivery point	via Bassini, 15		
City	Milano		
Postal code	20133		
Country	Italy		
Electronic mail	pepe.m@irea.cnr.it		
OnLine resource	www.aware-eu.info		
Role	distributor		
Name	GeoTIFF		
Version	6		
WMS Map Server	http://telea.mi.cnr.it/cgi-bin/mapserv?map=/opt/aware/private/aware1/modis.map; OGC:WMS-1.1.1-http-get-map; MOD101A-2007280-32N-TIF		
Spatial representation info			
No. of dimensions	2		
Dimension name	Column	Row	
Dimension size	2478	2300	
Resolution	500	500	
Cell geometry	area		
Reference system info			
Code	EPSG:32632 - WGS84 / UTM zone 32N		
Data quality info			
Hierarchy level	Data set		
Statement	The Terra product, MOD10A1, is validated (stage 2); the Aqua product, MYD10A1, is provisional.		
Metadata			
File identifier	{EAE1281C-F18A-4F5E-9809-C2CF82343A03}_100029_en		
Character set	utf8		
Hierarchy level	Data set		
Date stamp	2007-12-20T09:12:41		
Metadata standard	ISO 19115:2003/19139		
Standard version	1.0		
Access constraints	intellectualPropertyRights		
Use constraints	intellectualPropertyRights		
Content info			
[gco:RecordType]	Key: 0=missing data, 1=no decision, 11=night, 25=no snow, 37=lake, 39=ocean, 50=cloud,		

	100=lake ice, 200=snow, 254=detector saturated, 255=fill
Content type	Thematic Classification
Imaging condition	Cloud
Code	Passed
Cloud cover %	30

- *classification system* by which the data are organized in one or more hierarchies within a Thesaurus;
- *administrative data* concerning property and rights of the data, costs, data payment and purchase modes and supports on which they are distributed;
- *information about metadata* including the date in which they were produced, the date of the last test on the metadata, the date of the last update, the date of the next test, the spatial reference system used to specify the geodata set‘s plane extension.

As far as *data quality* is concerned, it is suggested to always edit a report useful for a future data utilization that can differ from the original one. If a report does not come with the acquired data, it is advisable to ask for the following questions about the data:

- To what period do they refer?
- Where do they come from?
- How were they collected?
- What software was used to produce them?
- What area do they cover?
- At what scale have they been digitized?
- What is the reference system, projection, datum?
- What are frequency and density of the observations?
- What is the attribute and position accuracy?
- Are they logic?
- Are they relevant for the considered application?
- What is their current format and what was their original one?
- Have they ever been tested and how?
- Why have they been produced and by whom?

The right answer to these questions certificates data quality defining their actual usefulness.

9.12 Geographical Information Systems Distribution on the Web

The GIS data are a patrimony to be protected by the administrative authority in charge of managing the system; it should even coordinate the required data updates

and organize them into categories longing for a complete control to the accesses of the system itself.

To query or consult a Geographical Information System, the authorized operators have to be provided with the appropriate software (the same used during the GIS designing/building or a compatible one). This implies the adoption of specific and suitable hardware configurations as well as liveware permitting to use the system. This is a limit to the spread and sharing of GIS resources. A strategy to overcome this problem consists in the so-called WebGIS.

The development of tools aimed at Web publication of maps and geographical data is greatly increasing: they can be considered as software designed for Web servers that allow distribution on the Internet of geographic data stored in the server; the server acts as a processor devoted to the database spreading, even according to very complex network architectures.

The World Wide Web is becoming a useful tool for spatial data sharing and is especially efficient at quickly transmitting territorial information, for example in civil emergency situations.

9.12.1 Requirements and Purposes of a WebGIS

Employing Internet technologies for access to GIS data is easy and cheap; a connection to the network and an appropriate browser equipped with opportune plug-ins are required.

Geographical Information System diffusion via network (GeoService) has to satisfy some fundamental requirements:

- the possibility of extracing descriptive information and data from the database associated with the available maps;
- map data visualization capability;
- the possibility of correlating geometrical features to the related alpha-numerical data;
- complete integration in hypertextual documents possibly manageable by HTML instruments.

Security is a keyword in the use of the network to distribute Geographical Information System. The original data must be protected from unauthorized access and editing operations potentially performed by remote users.

Some of the existing software tools allow the definition of different access levels; the GIS administrator can decide how data can be accessed by different user categories. As far as the graphical visualization of data is concerned, they can be supplied to the client in raster (more common) or vector formats; this operation is carried out by *extended tools* provided together with the main software. Maps visualized on the client are usually not georeferenced products even if they maintain the correct scale factor. This strategy prevents the unauthorized stealing of data: in fact, although they can be copied into the PC cache, CAD or GIS software cannot process them as georeferenced data. This way preservation of the original data's economic

value is granted. These softwares have been conceived just as tools for GIS contents visualization, but not as a tools of free distribution of the data.

It is worth remembering that data acquisition is one of the highest costs in the definition of a GIS and hence it is important to assure the highest visibility by the highest number of users as possible, always preserving the data property.

Based on the typology of end users, WebGIS can be divided into three main classes:

- intranet dissemination of data according to different levels of privileges;
- public, accessible by any Internet user and devoted to the distribution of information available to users' needs;
- digital catalogues containing data that can be bought by other organizations.

From the user's point of view, the data research and selection is carried out by a pre-defined interactive net surfing, which exploits the connections among the data, or directly a search engine able to formalize specific queries.

The quick convergence between the Geographical Information Systems, the Satellite Positioning Systems and the most advanced informatics and telecommunication technologies produce services both in the public and private fields, for example for the developing of localization services. These are supplied by the ASP (*Application Service Provider*): ASP provides a service for distribution, managements and updating of software applications working on the Internet and hosted in an IDC (*Internet Data Centre*). Users' access type is permitted according to a contract that does not require licences purchase, but periodical fixed, or indexed by a use parameter (pay and use) fees.

The ASP approach allows the administrators/companies to immediately activate GIS functions in their information systems, avoiding designing, building, updating and maintaining their own DBMS with spatial information.

Data collection, management and maintenance are a relevant part in a GIS.

Federated systems avoid data duplications leaving data updating in charge of the same institution that attended to their collection.

9.12.2 Federated and Distributed Systems

An important tendency in building complex GIS is the necessity to develop systems where the geographic database is split into different remote locations (Fig. 9.39).

A system can be

- *integrated*, if the database is fed with data coming from different sources accessible by remote users;
- *distributed*, if the system is centrally managed and physically decentralized;
- *federated*, if both the management and the physical location are centralized.

In distributed databases, the user ignores the design of the system he/she is using, where data are and how they are represented; on the other hand, the system has to guarantee the access defining a sharing scheme that can vary in time.

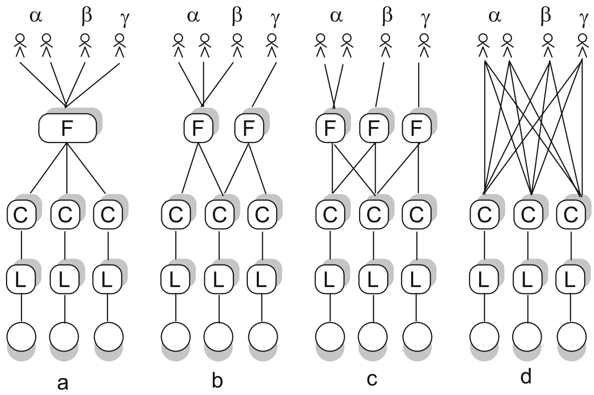


Fig. 9.39 Architectures for a distributed databases system: (**a**) Global Schema, (**b**) Multiple Integrated Schema, (**c**) Federated Schema, (**d**) Multi-database Language. The users at the sites α, β, γ access the data from three sources. A Local Schema (L) is the conceptual schema of a component database system and can be expressed in different data models (such as the component database systems). For each local schema, there is a corresponding Component Schema (C). Component (C) and Federated (F) Schemas are expressed by a Canonical Data Model (CDM). A Federated Schema is the integration of multiple Component Schemas. To each C is associated an L, namely the conceptual schema of a Component Database System (CDS)

9.12.3 Structure of GIS Diffusion Systems on the Web

Software for Web geographic data publication runs inside an appropriate processor (Web server) connected to the Web.

The WebGIS basic scheme can have different variants on the client side as well as on the server side (Fig. 9.40).

The user, surfing the Web as a client operator, accesses the system through *Web browsers* able to manage HTML documents and some standard format images: *Tag Image File Format* (TIFF), *Graphic Interchange Format* (GIF), *Joint Photographic Experts Group*(JPEG), *Portable Network Graphics* (PNG). In order to operate even with other data formats (e.g. vector data, audio or video files), the browser functions have to be extended according to one of the following strategies:

- plug-ins: additional routines running on the client, downloadable from the Web. In general software version, problems can occur ; thus they have to be constantly updated;
- Active-X controls: software extensions executing some tasks and able to communicate with other programs and play the same role as the plug-ins; the main difference is that the Active-X controls, besides extending the browser functions, can be used by programs or applications supporting the OLE (*Object Linking Embedding*) standard;

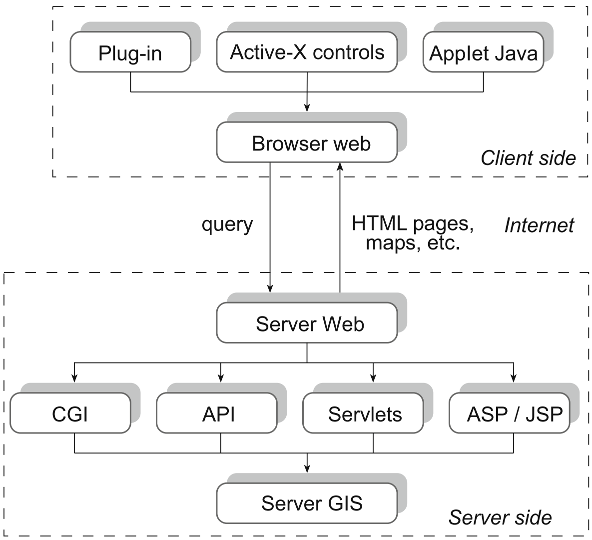


Fig. 9.40 WebGIS structure and possible link systems between the client and a GIS on line. Plug-in, Active-X controls. Applet Java: programs that enrich the browser; CGI: Common Gateway Interface; API: Application Program Interface; ASP: Active Server Pages; JSP: Java Server Pages

- Applets Java are routines executed by the browser onboard of the client; an applet is an executable that, for security reasons, can only access the server which has been downloaded from, and it is independent from the hardware platform.

In WebGIS the Web browser visualizes HTML pages, containing the GIS server content, dynamically; this can occur in different ways:

- *CGI Program* (*Common Gateway Interface*) able to recall a program required to establish a connection to a server database supplying the information to be inserted in the HTML page; CGI technology is independent from the adopted programming language (Fig. 9.41);
- *Web server− API* (*Application Program Interface*), programmable interface resident on the server; it permits partially optimization of the server memory, since there processes do not run independently, as in the previous case (Fig. 9.42);
- *Active Server Pages* (ASP), server-side programs operating in the Microsoft environment, enabling the development of dynamical and interactive applications for a Web server;
- *Java Server Pages* (JSP), technology similar to the ASP, included in the HTML code; these applications are written in Java language, thus they can run onboard any operative system and server;
- *Java Servlet* is the answer to CGI programming through Java technology; it is a group of programs running on the server and producing HTML pages; the advantages are the higher efficiency and the higher portability.

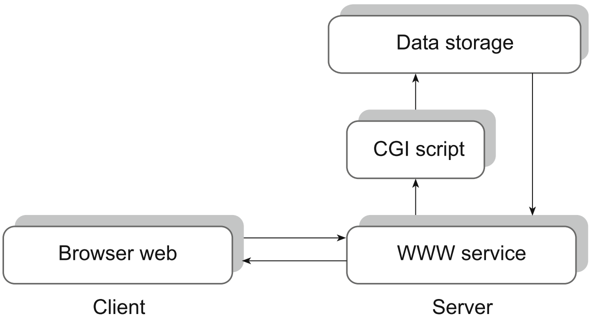


Fig. 9.41 Operability of CGI (Common Gateway Interface) applications

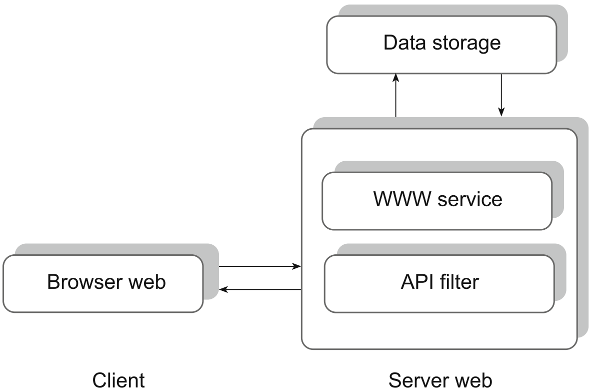


Fig. 9.42 Operability of API (Application Program Interface) applications

Online GIS support both the vector and the raster format.

The choice among different format types depends on the application to be developed, the operative infrastructure (raster format was conceived for the Web in TIFF, GIF, JPEG formats, while vector format needs plug-ins, and generally both require contacts with the server for each single client request) and to the adopted software.

9.12.4 Architecture of a Web-Oriented GIS

The component of a WebGIS conditions the architecture model, and they determine high levels of:

- *reliability*: capacity to provide a service, facing any breakdown;
- *service*: performance in terms of time of the system (answering time for user requests);
- *scalability*: adaptability to the visualization of arbitrary large documents;

- *security*: prevention against unauthorized users.

The definition of a structure is a hard task since, in the Web domain, the interaction between the different components is very high. For example, the connection speed affects the amount of transferred data, the data types affect the volumes to be transferred, the number of users affects both the connection and the system performance, the control of the access permissions affects the data transfer speed.

Thus, it is important, in the starting phase, to think globally and later detail the several issues in order to reduce the model complexity.

An architecture model presenting high complexity as well as strong generality characteristics is proposed in Fig. 9.43 and can be considerably simplified according to the specific application.

Within the global scheme, interconnected sub-systems are described at different detail levels in function of the requirements:

- *Data Repository*: indicates the set of storing tools (Disk Array, Tapes Libraries, CD Array, DVD, Flash Memory, etc.) containing the raw data whose integrity and coherence must be preserved;
- *Storage Area Network* (SAN): a set of very sophisticated hardware components, connected one to the other by optical fibre, together with the necessary software assuring a complete scalability (necessary for amounts of data that can be over the a Terabyte) and data transfer services that can reach a Gigabyte per second;
- *Application Servers* area: refers to the set of tools, included their software, which manage all of the applicative services. In the scheme, only some servers are

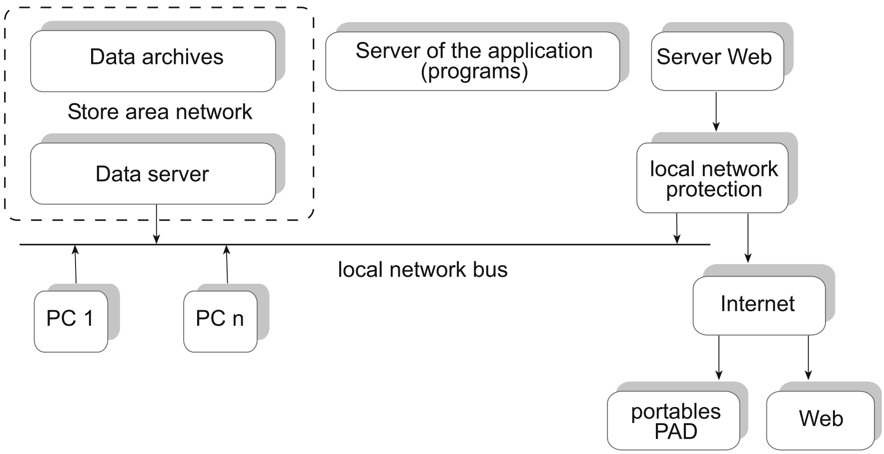


Fig. 9.43 Architectural model for data interchange on the Web. PDA: Personal Digital Assistant

shown (Terminal Server, Map Server) representing just some of the most common applications;

- *Security Server* area: refers to the set of tools, included their software, managing the system protection. The scheme includes the use of firewalls, specific tools for the control of the traffic, the monitoring of access, and separating the external from the internal network;
- *Web Server* area: refers to the set of tools, including their software, that allow the system to organize the information and to make them available online;
- *Network Communication* area: refers to the set of tools, included their software, representing the integrated communication structure aimed at transporting the data from one location to another. GIS applications innately process large amount of data and consequently the Web is kept under stress. Hence the design and the dimensioning of the system have to be accurate;
- *client* area: refers to the set of tools, including their software, aimed at producing requests to the server and obtaining the information from it. The model defines the high variability of these instruments, from the PC to the mobile phone. The Geographic Network is the set of geographic resources available in the Web activating a bi-directional link with the system.

9.12.5 Applicative Software

The main properties of software tools for distribution of maps, geographic data and GIS via Web are

- possibility of extracting, by query from a related database, information about a specific object;
- vector data management;
- graphic tools for visualizing maps: enlargement and dragging tools, representation scale management, etc.

The most innovative issue is the publication of *active maps* whose graphical elements can be activated by mouse movements on the screen, which recall feature attributes in the database and enable dynamic links.

9.12.6 Data Interoperability

Geographical Information Systems, to be effective and efficient for some applications, require to be accessible in real time and to interact with the operator/user even if no direct access to the physical system hosting and managing the application is active. The possibilities offered by the Web for the data spread allow transferring and processing of spatial information. For this reason, in recent years, several GIS producers started to sell online versions of their systems, allowing the users

to remotely access the GIS. The GIS role is generally of as a *distributor* of maps organized in thematic layers that can be overlaid and processed by some functions typical of the GIS. Geographical data are generally big (sometimes gigabytes), and their transfer and accessibility represent a complex problem.

The data interoperability issue is still more complex due to the quick development and increasing competition among GIS products, developed by many international companies to be used at global scale. Each producer, to gain wider and wider market sectors, tries to impose its own standards regarding data format, transfer protocol, etc. These are often not compatible with each other, and their re-formatting is very expensive.

With *interoperability* is intended the capacity or possibility to access different systems and spatial data, also organized with different software tools, recognizing and adapting different formats, spatial models, processing, applications without modifying their information content and quality.

Considering the increasing dissemination of terrestrial information and the favourable development of the market, in 1998 the *Open Geospatial Consortium* Inc.® (OGC) was constituted. This Consortium is a non-profit, international, voluntary consensus standards organization that is leading the development of standards for geospatial- and location-based services. It aims at facilitating interoperability and data sharing.

The main purpose of the OGC is to define a neutral data format not linked to any GIS software producers. One of the suggested solutions is the definition of a free language called *Geography Markup Language* (GML), specifically defined to codify, store and transfer geographic data, also via the Web.

Based on the XML (*eXtensible Markup Language*), it manages geographic elements geometry and attributes. If adopted this would enable the producers to share heterogeneous data sets and the users to transparently access them. Other words, it would simplify and standardize the operations in different sectors, from map production to data transformation, from spatial queries to geographic analysis, including the emerging applications in mobile systems.

9.12.7 XML Standard

EXtensible Markup Language (XML) standard, developed by a standardization consortium, the W3C (*World Wide Web Consortium*) allows representation, storage and exchange of data in an extendible and flexible way, according to an appropriate paradigm, depending on the data and on the application. It represents a tool to codify the data in textual mode and can combine a large variety of data types, such as texts, numerical, graphic, audio, images formats and maps, thus including the spatial information. XML technology nowadays is spread and included in some Web browsers. While the HTML defines the content together with its presentation, the XML separates them. This means that an XML file can be formalized in different ways according to the tool that will use it, to the user and his skill, etc. XML requires definition of the way a specific content must be represented; for this rea-

son XML transformation technologies have been developed, for a wide variety of devices, such as *Personal Digital Assistants* (PDAs) with wireless connections. Like HTML, XML is a marking language, which is a language using tags to characterize document contents. While in HTML, the tags are pre-defined independently from the domain of the application, in XML the tags have to be defined by the data producer and can vary according to the specific application. Both the meaning and the structure of the specific contents condition them. This characteristic induced the development of XML *dialects*, each one adapted to define different application fields' documents. An example is MathML, for Mathematics, or *Chemical Markup Language* (CML) for Chemistry. GML is the equivalent of XML for Geography.

For this reason, XML has to be considered a meta-language, namely a language for the definition of other languages or applications.

In XML the document structure is defined by a DTD (*Document Type Definition*) model or a specification according to a Schema; a DTD model or a Schema provide a formal description of the marking language created by the user: they describe a list of the admitted tags, their structural relations, their attributes and instructions (or restrictions). The latter are used by the validating analysers as rules to verify the elements, the attributes, the identities and the notes of XML documents defined according to that DTD or Schema.

As distinct from HTML, XML provides a single XML-element, a kind of capillary linking mechanism, and parts of it, with complex associations, as well as bi-directional ones.

9.12.8 Geography Markup Language (GML)

GML is a marking language belonging to the XML family, developed and maintained by OGC, used to represent geographic data and their meaning. Geographical data describe the world's geometry independently from the way it is visualized. The GML has its own structure to code spatial data and their associations; it can also be extended to fit a specific application context. In general, GML has the capability to combine and associate geospatial data with other types of data (expressed in XML or GML), giving a solution to the problem of the integration of heterogeneous data.

At its lower level, GML provides a method to describe the spatial elements and their attributes in a textual mode within an XML document. These documents are easy to interpret, and they can successfully formalize the relationships between both different object geometries and attributes, profiting from XML's high linkage capacities. GML is not a data presentation language, and thus has to be properly transformed to be presented to the user, generally in map or image format, but also in text or even vocal format.

GML implementation mainly provides a tool to transfer and store data in the Web domain. It provides a wide variety of spatial objects; it can keep separated the concept of data from that of their visualization; it can relate spatial elements with their attributes allowing the definition of spatial relationships (Fig. 9.44).

Geographic data codification by languages like GML aims at solving the emerging interoperability problems related to data representation.

```xml
<gmd:referenceSystemInfo>
    <gmd:MD_ReferenceSystem id="ID0002">
        <gmd:referenceSystemIdentifier>
            <gmd:RS_Identifier>
                <gmd:code>
                    <gco:CharacterString>EPSG:32632 - WGS84 / UTM zone 32N</gco:CharacterString>
                </gmd:code>
            </gmd:RS_Identifier>
        </gmd:referenceSystemIdentifier>
    </gmd:MD_ReferenceSystem>
</gmd:referenceSystemInfo>
<gmd:identificationInfo>
    <gmd:MD_DataIdentification>
        <gmd:citation>
            <gmd:CI_Citation>
                <gmd:title>
                    <gco:CharacterString>MOD10A1 Snow_Cover_Daily_Tile 2007281</gco:CharacterString>
                </gmd:title>
                <gmd:date>
                    <gmd:CI_Date>
                        <gmd:date>
                            <gco:Date>2007-10-13</gco:Date>
                        </gmd:date>
                        <gmd:dateType>
                            <gmd:CI_DateTypeCode codeList="http://metadata.dgiwg.org/codelistRegistry?
                            CI_DateTypeCode" codeListValue="creation"/>
                        </gmd:dateType>
                    </gmd:CI_Date>
                </gmd:date>
......................................................................................................................................
```

Fig. 9.44 Example of metadata in XML format (Standard ISO:19139) referring to Snow Covered Area daily map products created by NASA and distributed in an INSPIRE compliant Spatial Data Infrastructure (courtesy of AWARE and IDE-Universe Projects, supported by European Commission)

The relevant aspects, distinguished in advantages and disadvantages, characterizing the use of this technology are

- GML language advantages:
  - *separation between content and representation*: GML contains geographic data, their geometry, their characteristics and attributes, but it does not declare how they have to be represented. Representation can be done through the adoption of style sheets: in such a way it is possible to visualize the data in different graphic formats (i.e. SVG, VML, etc.) and, within the same format, and further customize the style (symbols, lines, areas representation);
  - *interoperability*: the GML is a standard textual format, without ownership, conceived to facilitate the sharing of heterogeneous data sets and to overcome the differences between the different GIS data formats;
- GML language disadvantages:
  - *incomplete standard*: the biggest problems currently affecting GML are mainly due to some gaps in the available descriptors; this makes, for example, the integration of data having a different reference systems quite complex;
  - *attributes management*: GML specifications regarding attribute management are quite general at the moment. If this can be seen as a positive aspect for the

language's expansibility, on the other hand it is an obstacle in the definition of applications aimed at managing these data.

9.12.9 Instruments for Graphical Representation

GML's primary goal is to define and exchange geographic data content. It can also be used to represent those data in map format, by the use of a rendering instrument able to interpret GML data. Thus GML objects have to be transformed into a format that can be opportunely interpreted by a graphical visualizer inside a browser. This requires the interpretation of the GML as graphical symbols. The *eXtensible Stylesheet Language Transformation* (XSLT), itself part of XML, can be the tool to build the presentation of data. It uses *eXtensible Stylesheet Language* (XSL), a language for the transformation of an XML document (e.g. GML) into another XML document (e.g. *Scalable Vector Graphics*, SVG) according to transformation rules specifically defined.

Generally this *graphical rendering* requires that GML data are transformed into an XML graphic format; during the transformation process, the data model can be changed and a filtering of the contents, aimed at extracting only what is necessary in a specific context, is possible. There are XML *dialects* that were created specifically to visualize graphical data (not necessarily geographic ones); some among the most well known are

- *Scalable Vector Graphics (SVG)*, for the definition of 2D vector graphics;
- *Vector Markup Language* (VML), (SVG competing proposed by Microsoft);
- VRML, for 3D vector graphics.

The graphical languages derived from XML have two important characteristics:

- *interactivity*: the document to be visualized is sensitive to some actions (like the mouse click) carried out by the operator producing specific actions as a reply to these events;
- *dynamics*: it is possible to create some animations by SMIL (Sy*nchronized Multimedia Integration Language*) animation language, developed by W3C.

9.12.10 Graphical Representation of Geographic Elements

Visualization languages for geographic documents, like GML files, have to represent geographic information not only in raster format, as traditionally happens on the Web, but also in vector format; moreover it is necessary to have the possibility to group the data represented in different layers with different content. Generally a digital map can be considered as an overlaying of different thematic vector layers, separately controllable; both GML and visualization languages offer this possibility.

Files defined by visualization languages like SVG and VML can be visualized in a Web browser, by installing specific plug-ins free. The basic functions offered to the user are

- map visualization;
- map visualization management: pan, zoom, symbols, colours, styles, tags, etc.;
- thematic layer management by hierarchical organization;
- data classification according to the attributes and subsequent graphic representation;
- thematic layer attributes description;
- selected elements description;
- data source visualization.

9.13 Spatial Data Infrastructure

The great request of GeoSpatial Data and their management induced several organizations worldwide to improve interoperability for data sharing.

The correct interoperable and interdisciplinary use of the *GeoSpatial Information* (GI) departs from the definition of a Spatial Data Infrastructure (SDI). For analogy, an SDI can be thought of as a transit infrastructure enabling the mobility of people and goods, including rules (i.e. distance travelled on roads), facilities (i.e. railway, channels) and organization (i.e. traffic management) overcoming the administrative boundaries within and among countries. In the same way, the GI must be based on concrete and lasting elements.

The *Spatial Data Infrastructures* (SDIs) are intended to allow the sharing, on the Web, of geocoded spatial data that are heterogeneous by definition, in terms of management system and development scale.

The concept of an SDI is based on the fundamental principle that systems and databases must conform to some basic requirements, thus creating a real infrastructure allowing the common use of data at any level at which they were generated.

In the field of GeoSpatial Information, there have been activated (or there are coming in the near future) programs that will influence research, enterprises and professional activities at International levels in the medium to long term, as the definition of Spatial Data Infrastructures at National, Continental and global level. Several initiatives are in progress to improve the knowledge of our planet in terms of GeoSpatial Information, and GEO and GEOSS programs are among them.

The focal point is to promote and orientate processes of geospatial harmonization in which different public, first, and private bodies and organisms adopt the same rules.

9.13.1 GSDI

Geographic Information is vital to take decisions at the local, regional, and global levels, and decision makers benefit from GI, with the infrastructures involved in the

process. Thus, the definition of shared guidelines aimed at regulating the basics of a Spatial Data Infrastructures is crucial.

These guidelines have been presented by Douglas (2004) in *The SDI Cookbook, Developing Spatial Data Infrastructures* – from this document are reported the following notes.

The SDI Implementation Guide or Cookbook, through the support of the Global Spatial Data Infrastructure community, supplies geographic information providers and users with the necessary background information to evaluate and implement existing components of SDI. It also facilitates participation within a growing (digital) geographic information community known as the Global Spatial Data Infrastructure (GSDI).

To enable builders of SDI to make use of and build on existing SDI components in a way which makes their endeavours compatible with the efforts of other SDI builders, this GSDI Cookbook identifies

- existing and emerging standards;
- open-source and commercial standards-based software solutions;
- supportive organizational strategies and policies;
- best practices.

Working within a common framework of standards and tools based on these standards also makes it possible to maximize the impact of the total available resources for SDI creation through future cooperation, e.g. we develop this, you develop that, and then we share.

At a global scale, the most prominent examples of formal SDI programs are on a national scale. Most of these are driven by the national or federal government (e.g. the NSDI in the USA, the SNIG in Portugal, Australia's ASDI, Malaysia's NaLIS, South Africa's NSIF, Colombia, or the GSDI Cookbook, multi-national INSPIRE Initiative in Europe), but there are exceptions, which have largely been driven by the private sector.

The Cookbook does provide a basic set of guiding principles that have been successful for establishing compatible Spatial Data Infrastructures and are supported by the Global Spatial Data Infrastructure to promote successful decision making for issues of local, regional and global significance. It cannot claim to provide all the answers to suit all variations that may exist among implementations of national spatial data infrastructures. The goal is to provide enough common guidance to allow adjacent SDIs to exchange information easily through the adoption of common principles, standards and protocols.

The development of the Cookbook is envisaged as a means to clarify the SDI definition and to share current experiences in building SDI implementations that are compatible at many scales of endeavour on the following arguments:

- geospatial data development: building data for multiple uses;
- metadata: describing geospatial data;
- geospatial data catalogue: making data discoverable;
- geospatial data visualization: online mapping;

- geospatial data access and delivery: open access to data;
- outreach and capacity building: creating a community.

9.13.2 Infrastructure for Spatial Information in the European Community – INSPIRE

The realization of an SDI for Europe is the goal of the European initiative INSPIRE.

The Directive 2007/2/EC of the European Parliament and of the Council of Europe on March 14, 2007 establishing an *Infrastructure for Spatial Information in the European Community* (INSPIRE) was published in the official Journal on April 25, 2007. The INSPIRE Directive entered into force on May 15, 2007.

The Article 1 of the Directive clearly focuses the objective:

1. *The purpose of this Directive is to lay down general rules aimed at the establishment of the Infrastructure for Spatial Information in the European Community (INSPIRE) for the purposes of Community environmental policies and policies or activities which may have an impact on the environment.*
2. *INSPIRE shall build upon infrastructures for spatial information established and operated by the Member States.*

The solution suggested by INSPIRE proposes a compromise, without forcing societies and users, to adopt a common data definition format.

The general situation on spatial information in Europe has generated in the time the fragmentation of data sets and sources, gaps in availability, lack of harmonization between data sets at different geographical scales and duplication of information collection. These problems make it difficult to identify, access and use available data from spread sources.

Fortunately, awareness is growing at national and at EU level about the need for quality georeferenced information to support understanding of the complexity and interactions between human activities and environmental pressures and impacts. The INSPIRE initiative is therefore timely and relevant but a major challenge given the general situation outlined above and the many stakeholder interests to be addressed.

INSPIRE is complementary to related policy initiatives, such as the Commission proposal for a Directive on the re-use and commercial exploitation of Public Sector Information.

The initiative intends to trigger the creation of a European spatial information infrastructure that delivers to the users integrated spatial information services. These services should allow the users to identify and access spatial or geographical information from a wide range of sources, from the local level to the global level, in an interoperable way for a variety of uses. The target users of INSPIRE include policy-makers, planners and managers at European, national and local level and the citizens and their organizations. Possible services are the visualization of information layers, overlay of information from different sources, spatial and temporal analysis, etc.

In Europe several countries have begun the process of realization of the Spatial Data Infrastructure (SDI) based on the following general principles:

- data should be collected once and maintained at the level where this can be done most effectively;
- it should be possible to combine seamlessly spatial data from different sources and share it between many users and applications;
- spatial data should be collected at one level of government and shared between all levels;
- spatial data needed for good governance should be available on conditions that do not restrict its extensive use;
- it should be easy to discover which spatial data is available, to evaluate its fitness for purpose and to know which conditions apply for its use.

9.13.3 GEO and GEOSS

Earth Observation systems consist of measurements of air, water and land made on the ground, from the air, or from space. Historically observed in isolation, the current effort is to look at these elements together and to study their interactions.

Producing and managing better information about the environment has become a top priority for nations around the world. In July 2003, the Earth Observation Summit brought together 33 nations plus the European Commission and many International Organizations to adopt a declaration that signified a political commitment towards the development of a comprehensive, coordinated and sustained *Earth Observation System* (EOS) to collect and disseminate improved data, information and models to stakeholders and decision makers.

An ad hoc group of senior political officials from all participating countries and organizations, named the *Group on Earth Observations* (GEO), has been formed to undertake this global effort. GEO was charged to develop a *Framework Document* plus a more comprehensive report to describe how the collective effort could be organized to continuously monitor the state of the environment, increase understanding of dynamic Earth processes, and enhance forecasts on environmental conditions. Furthermore, it was to address potential societal benefits if timely, high-quality, and long-term data and models were available to aid decision makers at every level: intergovernmental organizations, local government and individuals.

On February 16, 2005, member countries of the GEO agreed to a 10-year implementation plan for a *Global Earth Observation System of Systems* known as GEOSS. The GEOSS project aims to address nations involved to produce and manage their information in a way that benefits the environment as well as humanity.

GEOSS is envisioned as a large national and international cooperative effort to bring together existing and new hardware and software, making it all compatible in order to supply data and information at no cost. The developed nations have a unique role in developing and maintaining the system, collecting data, enhancing data distribution, and providing models to help all of the world's nations.

Outcomes and benefits of a global information system will include

- disaster reduction;
- integrated water resource management;
- ocean and marine resource monitoring and management;
- weather and air quality monitoring, forecasting and advisories;
- biodiversity conservation;
- sustainable land use and management;
- public understanding of environmental factors affecting human health and well-being;
- better development of energy resources;
- adaptation to climate variability and change.

9.13.4 Global Monitoring for Environment and Security, GMES

The European view of GEOSS is *Global Monitoring for Environment and Security* (GMES), which focuses on integrating the Systems of Systems with a chain of services calibrated on the necessity of Europe and its inhabitants.

The GMES initiative is, with the Galileo programme, the main element of the European Space strategic development.

The GMES initiative started in 1998 to bring together both users and suppliers of information derived from satellite observation. GMES has become a key element in the European space strategy, oriented towards strengthening the Community's ability to acquire high-quality information derived from space, atmospheric, terrestrial and marine observations integrated with geographical and socio-economic data.

Following endorsement by the EU Council, the Commission and the European Space Agency (ESA) are implementing the action plan on GMES developed in 2001.

The requirements that drive the implementation of GMES include openness, federated architecture, simplicity of architecture, scalability, dependability, user-friendliness, data security, quality of service and global ubiquity of access.

The European Commission's Directorate General *Joint Research Centre* (JRC) makes a substantial contribution to the GMES initiative through a wide range of capabilities. It has led the way in the establishment and development of databases of European geospatial information, successfully integrating information from satellites into databases as well as a range of other research and monitoring activities.

GMES consists of four main components as shown (Fig. 9.45):

- *space component*, which includes development and operations of satellites and ground segments infrastructures providing the required data streams;
- *in situ component*, which includes development and operation of all ground-based and airborne data gathering networks. This also includes socio-economic data such as demographic data, industrial output data and epidemiological data;

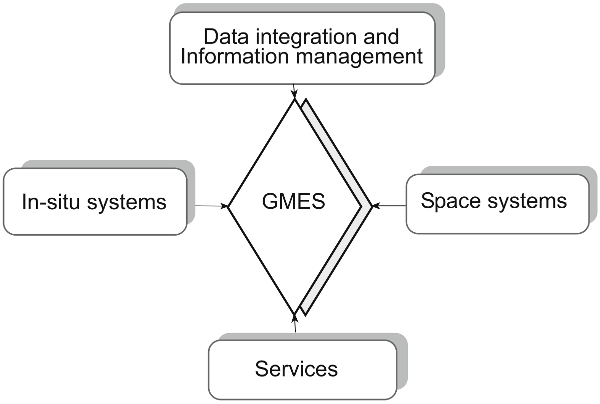


Fig. 9.45 The GMES (*Global Monitoring for Environment and Security*) capacity as represented by the *GMES diamond*, in the Communication from the Commission to the European Parliament and the Council: Establishing a GMES capacity by 2008 – Action Plan (2004–2008)

- *data harmonization* and standardization, including infrastructure for data assimilation, standardized processing, data fusion, modelling and forecasting and service support infrastructure such as data access, catalogue, archive and customer support facilities;
- *user-oriented services*, including data procurement, information generation and delivery, product quality assurance and monitoring compliance with user domain working practices, standards and legislative requirements.

These four elements are interconnected, following a modular approach, to serve the GMES general objectives.

9.14 Summary

Several authors (Burrough, 1986; Cowen, 1988; Aronoff, 1989) have defined Geographical Information Systems. Burrogh (1997) as *a powerful set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world for a particular set of purposes*.

The geographical feature, or ground element, is characterized by the positional and the descriptive components. According to this concept, GIS is a complex system able to manage the geographical position of objects, their attributes, spatial and attribute relationships and the time factor of dynamic phenomena. The attribute can be metrical, graphical, descriptive and complex. The spatial relationships can be topological, directional and of proximity.

The GIS components are hardware, software, data and liveware or organization context. Geographical data components are analysed and managed by *DataBase Management Systems* (DBMS). DBMS are described in Chapter 5. The spatial data model can be vector or raster, with possibility of integration of the two.

In GIS, methods of spatial data analyses are performed. The spatial relationships among the features and their attributes and the persistent link with their geometry make the GIS a tool able to simulate the real world. These methods are the spatial analysis of data, attribute analysis and their integrated combination.

The Earth's surface can be described in a *complete*continuous *representation* using different models for the visualization. The Digital Terrain Model (DTM) and its equivalent Digital Surface Model (DSM) and Digital Elevation Model (DEM) are described. In the processing flow, the altimetric data acquisition performs an important role. Methodologies and techniques of 3D points acquisition are mentioned and analysed: ground survey, digitization of existing maps, aerial and satellite photogrammetric methods (Chapter 3), radar interferometry (Chapter 4) and Airborne Laser Scanning (Chapter 6). The structural model for 3D representation, generation and interpolation is also discussed.

GIS are in continuous evolution and *object-oriented data modelling* is suitable for spatial applications to describe the elements in the real world and the relative properties as objects, in order to support the processing of complex geometric entities.

An important evolution of GIS is the Decision Support Systems (DSS). They host sophisticated processing tools, and they are conceived to answer the question: *what is going to happen if...?*. Another trend which might improve GIS spread, even and especially at a local level, is the Virtual GIS (VGIS), which needs knowledge-based systems, or Expert Systems.

Data management and the importation of new data into GIS and DSS often suffer from errors that can sometimes be forecast. Terminologies as accuracy, precision, data quality, tolerance and error are introduced. The possible types and sources of error are described, and their influence in the processing chain is discussed. Errors can be introduced into the system by propagation or cascade.

Data quality issue and metadata are fundamental in GIS and DSS. *Metadata* certify the global quality of data sets. A detailed example of metadata is reported.

The development of tools aimed at Web publication of maps and geographical data is greatly increasing. A strategy to share protected data sets is the WebGIS. Purposes and requirements for WebGIS planning are presented.

A topic of large interest is data interoperablility. The great request of GeoSpatial Data and their management induced several organizations worldwide to improve interoperability for data sharing.

The efforts to overcome the problem are introduced, included a description of the *EXtensible Markup Language* (XML) standard and its derivation Geography Markup Language (GML).

The correct interoperable and interdisciplinary use of the *GeoSpatial Information* (GI) departs from the definition of a Spatial Data Infrastructure (SDI). The Global Spatial Data Infrastructure (GSDI) is the worldwide SDI. The European Community has defined the INSPIRE initiative with the Directive 2007/2/EC of the European Parliament and of the Council of Europe of March 14, 2007.

The *Group on Earth Observations* (GEO) and the *Global Earth Observation System of Systems* known as GEOSS play a fundamental role in Earth Observation

management. The activities concern on how collective effort could be organized to continuously monitor the state of the environment, increase understanding of dynamic Earth processes and enhance forecasts on environmental conditions. The European branch of GEOSS is *Global Monitoring for Environment and Security* (GMES).

Further Reading

Aronoff S., 1989, Geographic Information Systems: a Management Perspective. WDL Publications, Ottawa, Canada.
Bernhardsen T., 1999, Geographic Information systems: An Introduction. John Wiley and Sons, New York, pp. 448.
Burrough P.A., McDonnell R.A., 2000, Principles of Geographical Information Systems, Spatial Information Systems and Geostatistics. Claredon Press, Oxford, pp. 306.
Cercone N., McCalla G. (Eds.), 1987, The Knowledge Frontier: Essays in the Representation of Knowledge. Symbolic Computation Series. Springer-Verlag, New York, pp. 552, ISBN 3-540-96557-2.
Clarke K.C., 2001, Getting Started with Geographic Information Systems. Prentice Hall, New Jersey, USA, pp. 352.
DorlingD., Fairbairn D., 1997, Mapping: Ways of Representing the World. Prentice Hall, Pearson Education, Harlow, England.
Douglas D.N., 2004, Developing Spatial Data Infrastructures: The SDI Cookbook, Technical Working Group, GSDI, Version 2.0, January 25, 2004.
Hearnshaw H.M., Unwin D.J., 1996, Visualization in Geographical Information Systems, John Wiley and Sons, New York.
Commission of the European Communities, 2007, INSPIRE Directive, 2007/2/EC of the European Parliament and of the Council of 14 March 2007 Establishing an Infrastructure for Spatial Information in the European Community. Official Journal of the European Union, 50, 25 April, ISSN 1725-2555, L 108.
Jones C., 1997, Geographical Information Systems and Computer Cartography. Pearson Education, Asia, Singapore, pp. 319.

Bibliography

Albrecht J., 1995a, Virtual Geographic Information System (VGIS), IEEE-HICSS 1995, Maui, Hawaii, January, pp. 141–150.
Albrecht J., 1995b, Semantic Net of Universal Elementary GIS Functions, Auto-Carto 12, Charlotte, NC, 27 Feb–2 Mar 1995, pp. 235–244.
Axelsson P., 1999, Processing of laser scanner data - algorithms and applications. ISPRS. Journal of Photogrammetry and Remote Sensing, 54: 138–147.
Antenucci J.C., Brown K., Croswell P.L., Kevany M.J., Archer H., 1991, Geographic Information Systems: A Guide to the Technology. Chapman and Hall, New York Vol. 10, pp. 211–236.
Aronoff S., 1989, Geographic Information Systems: A Management Perspective. WDL Publications, Ottawa Canada.
Baltsavias E.P., 1999b, Airborne laser scanning: basic relations and formulas. ISPRS. Journal of Photogrammetry and Remote Sensing, 54: 199–214.
Baltsavias E.P., 1999a, A comparison between photogrammetry and laser scanning. ISPRS. Journal of Photogrammetry and Remote Sensing, 54: 83–94.

Binaghi E., Cerrani I., Montesano M.G., Rampini A., 1996, A hybrid fuzzy expert system shell. In: Fuzzy Logic and Neural Network Handbook, C.H. Chen (Eds.). McGraw-Hill Professional Publishing Group New York .

Binaghi E., 1990, A fuzzy logic inference model for a rule-based system in medical diagnosis. International Journal of Knowledge Engineering: Expert Systems, Learned Information, 7(3): 134–141.

Binaghi E., Della Ventura A., Rampini A., Schettini R., 1993, Fuzzy reasoning approach to similarity evaluation in image analysis. International Journal of Intelligent Systems, 8(7): 749–769.

Brachman R.J., Levesque H.J. (Eds.), 1985, Readings in Knowledge Representation, Morgan Kaufmann, San Mateo, California.

Brivio P.A., Marini D., 1993, A fractal method for digital elevation model construction and its application to a mountain region. Computer Graphics Forum, 12(5): 297–309.

Carrara P., Fortunati L., Fresta G., Gomarasca M.A., Piazza Bonati L., Poggioli D., 2004, A methodological approach to the development of applications in a SDI environment. 7th AGILE Conference on Geographic Information Science.

Carrara P., Fresta G., Gomarasca M.A., Pasi G., Poggioli D., Rampini A., 2003, A reversible web architecture for the management of geographic information. 6th AGILE Conference on Geographic Information Science, Lyon (F), 24–26 April, pp. 161–168.

Chen C.H. (Ed.), 1996, Fuzzy Logic and Neural Network Handbook. McGraw-Hill Professional Publishing Group, New York.

Cercone N., McCalla G., 1983, The Knowledge Frontier, Springer and Verlag, Berlin.

Cortellessa C.M., 1995, Breve Introduzione al GIS, Supplemento a Mondo Autocad, n. 5.

Cover R., 2001, The XML Cover Page – Geography Markup Language (GML).

Cowen D.J., 1988, GIS versus CAD versus DBMS: What are the Differences? Photogrammetric Engineering and Remote Sensing, 54: 1551–1554.

Cowen D.J. et al., 1995, The Design and Implementation of an Integrated Geographic Information System for Environmental Applications. Photogrammetric Engineering and Remote Sensing, 61(11): 1393–1404.

Dikau R., 1989, The application of a digital relief model to landform analysis in geomorphology. In: Three Dimensional Applications in Geographical Information Systems, J. Raper (Eds.). Taylor and Francis, London, New York, Philadelphia, pp. 51–78.

Duda R.O. Hart P.E., 1973, Pattern Classification and Scene Analysis. John Wiley and Sons, New York.

Ebner H., Reinhardt W., Hößler R., 1987, Generation, management and utilization of high fidelity digital terrain models. International Archives of the Photogrammatic and Remote Sensing, 27: 556–566.

Ehlers M., (1995a), Integrating remote sensing and GIS for environmental monitoring and modelling: where are we? In: Manfred Ehlers (Hrsg.), Aktuelle Forschungsberichte der Abteilung Geographische Informationssysteme/Fernerkunung. Vechta, Teil B, pp. 56–65.

Ehlers M., 1995b, Integration of remote sensing with geographic information systems: a necessary evolution. Photogrammetric Engineering and Remote Sensing, 55(11): 1619–1627.

Elliotte Rusty H., 2001, XML Bible – XSL Transformations, pub-HUNGRY-MINDS, xxxiii + pp. 1206, ISBN: 0 7645-4760-7.

Fisher P.F., Linderberg R.E., 1989, On distinctions among cartography, remote sensing and GIS. Photogrammetric Engineering and Remote Sensing, 55: 1431–1434.

Harrington S., 1983, Computer Graphics: A Programming Approach. McGraw-Hill. Book Company, International Editions, New York.

Healey R.G., 1991, Database management system. In: Geographical Information Systems: Principles and Applications, Maguire D.J., Goodchild M.F., Rhind D.W. (Eds.). Longman, London, Vol. 1, pp. 251–267.

Hinton J.C., 1996, GIS and remote sensing integration for environmental applications. International Journal of Geographic Information Systems, 10(7): 877–890.

Hoppe H., 1994, Surface Reconstruction from Unorganized Points. PhD Thesis. University of Washington, p. 116.

Jackson M.J., 1987, Digital cartography, image analysis, and remote sensing, towards an integrated approach. Interdisciplinary Science Reviews, 12(1): 33–40.

Jain R., Kasturi R., Schunck B., 1995, Machine vision. McGraw-Hill Book Company, International Editions, New York. p. 549.

Kandel A., Langholz G. (Eds.), 1992, Hybrid Architectures for Intelligent Systems. CRC Press, Boca Raton, Florida, USA.

Korpela I., 2000, 3D Data Capture for DEM/DTM/DSM Production - An Introduction to Photogrammetric Methods and Ranging Laser, Course Y196. Department of Forest Resource Management, University of Helsinki, Helsinki.

Koeln G.T., Cowardin L.M., Strong L.L., 1994, Geographical Information Systems, In: Bookhout Ed. The Wildlife Society, Bethesda, p. 540.

Kraus K., 1993, Photogrammetry. Vol. 1. Fundamentals and Standard Processes. Dummler, /Bonn, pp. 397.

Kraus K., 1995, From Digital Elevation Model to Topographic Information System, Photogrammetric Week'95, Fritsch/Hobbier Eds, Wichmann, pp. 277–285.

Kraus K., Pfeifer N., 1998, Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS. Journal of Photogrammetry and Remote Sensing, 53: 193–203.

Kufoniyi O., 1995, An introduction to object-oriented data structures. ITC Journal, 1: 1–7.

Lake R., 2001a, GML 2.0 – Enabling the Geo-spatial Web. Galdos Systems Inc. Vancouver, Bc Canada

Lake R., 2001b, Making Maps with Geography Markup Language (GML). Galdos Systems Inc.

Lake R., Burggraf D., Trinic M., Rae L., 2004. GML Geography Mark-Up Language. John Wiley and Sons, Southern Gate, Chichester, West Sussex.

Laurini R., Thomson D., 1992, Fundamental of Spatial Information Systems. The Apic Series, San Diego, Ca.

Maguire D.J., Goodchild M.F., Rhind D.W., 1991a, Geographic Information Systems, Vol. 1 and 2: Principles and Volume 2: Applications. John Wiley and Sons, New York, USA.

Maguire D.J., 1991b, An overview and definition of GIS. In: Geographical Information Systems: Principles and Applications, D.J. Maguire, M.F. Goodchild, D.W. Rhind (Eds.). Longman, London, Vol. I, pp. 9–20.

Marangoz A.M., Karakıs S. Akcın H., 2007, Object-based automatic classification of urban open green areas using high resolution QuickBird imagery and integration to GIS. Proceedings of the 27th EARSeL Symposium, GeoInformation in Europe, M.A. Gomarasca (ed.). Millpress, The Netherlands, pp. 125–134, ISBN 9789059660618.

Marble D.F., Peuquet D.J., 1990, Introductory Readings in GIS. Taylor and Francis, London, New York, Philadelphia.

Meijerink A.M., Valenzuela C.R., Stewart A., 1988, The integrated land and watershed management information system. ITC Publication, 7: 1–115.

Meijerink A.M.J., De Brouwer H.A.M., Mannaerts C.M., Valenzuela C.R., 1994, Introduction to Use of Geographic Information Systems for Practical Hydrology. UNESCO-ITC Publication Number 23, Enschede, The Netherlands.

Michalsky R.S., 1980, Pattern recognition as rule-guided inductive inference. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-2: 349–360.

Michalsky R.S., 1994, Inferential Theory of Learning: Developing Foundations for Multistrategy Learning, Machine Learning - A Multistrategy Approach, R. Michalski, G. Tecuci (Eds.). Morgan Kaufmann Publishers, San Francisco, Vol. 4.

Migliaccio F. 1998, Appunti delle lezioni del corso di Cartografi a tematica e automatica. CLUP, Milano

Neale I.M., 1988, First generation expert systems: a review of knowledge acquisition methodologies. Knowledge Engineering Review, 3(2): 105–137.

Pedrycz W., 1990, Fuzzy sets in pattern recognition: methodology and methods. Pattern Recognition, 23: 121–146.

Peuquet D.J., Marble D.F., 1990, Introductory Readings in Geographical Information Systems. Taylor and Francis, London, New York, Philadelphia, pp. 1–371.

Quin L., 2000, Open Source XML Database Toolkit, Wiley and Sons, New York.

Rott H., Nagler T., Malcher P., Mùller F., 2007, A satellite-based information system for glacier monitoring and modeling. Proceedings of the 27th EARSeL Symposium, GeoInformation in Europe, M.A. Gomarasca (Ed.). Millpress, The Netherlands, pp. 395–402, ISBN 9789059660618.

Rumelhart H., Hinton G.E., Williams R.J., 1986, Learning internal representation by error propagation. In: Parallel Distributed Processing, H. Rumelhart, J.L. McClelland (Eds.). MIT Press, Cambridge, USA, pp. 318–362.

Smets P., Mamdani E., Dubois D., Prade H., 1988, Non Standard Logics for Automated Reasoning. Academic Press, New York.

Sonka M., Hlavac V., Boyle R., 1999, Image Processing, Analysis and Machine Vision, 2nd ed. PWS Publishing, Pacific Grove, California p. 770.

Star J., Estes J., 1990, Geographical Information Systems: An Introduction. Prentice Hall, Englewood Cliffs.

Strobl P., Reithmaier L., Soille P., Mehl W., Bielski C., 2007, Assembly of a remote sensed reference image database of Europe at 25m resolution. Proceedings of the 27th EARSeL Symposium, GeoInformation in Europe, M.A. Gomarasca (Ed.). Millpress, The Netherlands, pp. 515–522, ISBN 9789059660618.

Tuffe E.R., 1990, Envisioning Information. Graphic Press, Cheshire, Connecticut.

Van Oosterom P., Zlatanova S., 2008, Creating Spatial Information Infrastructures, Towards the Spatial Semantic Web. CRC Press, Taylor and Francis Group, Boca Raton, London, New York, Leiden n. 70681, ISBN: 978-1-4200-7068-2.

Whiteside A., Bobbitt J., Nicolai R., 2001, Recommended Definition Data for Coordinate References Systems and Coordinate Transformations, OpenGIS® Consortium.

Wilkinson G.G., 1996, A review of current issues in the integration of GIS and remote sensing data. International Journal of Geographical Information System, 10(1): 85–101.

Yager R.R., Filev D.P., 1995, Including probabilistic uncertainty in fuzzy logic controller modeling using Dempster-Shafer Theory. IEEE Transactions on System, Man and Cybernetic, 25(8): 1221–1230.

Yager R.R., 1995, A unified approach to aggregation based upon MOM and MAM operators. International Journal of Intelligent Systems, 10: 809–855.

Zadeh L.A., 1965, Fuzzy Sets, Information and Control, 8: 338–353.

Zadeh L.A., 1973, Fuzzy Pattern Recognition. Academic Press, London.

Zadeh L.A., 1977, Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1: 3–28.

Zadeh L.A., 1981, PRUF - a meaning representation language for natural languages. In Fuzzy Reasoning and its Applications, E.H. Mamdani, B.R. Gaines B.R. (Eds.). Academic Press, London, pp. 1–58.

Chapter 10
Land Use/Land Cover Classification Systems

The final objective of a geomatics process concerning Earth observation is the production of thematic maps. How to obtain thematic results is described in Chapter 8. With the use of informatics instruments, the possibilities to produce thematic maps are extensive. The problem is the quality and the assessment of the reliability of these maps. Few rules are fixed in image classification, and one of the most important concerns legend definition. Due to the large variability of possible terrestrial classes in the world, each map producer organizes its own legend independently. The absence of a recognized common methodology is conducive to the production of a very large variety of thematic maps. The *queen* of the thematic maps is the land use/land cover map. But also the definition of *land use* and *land cover* introduces misunderstanding.

There are cases where some rules are fixed to harmonize the legends. The first is Anderson et al. (1972) who have defined the basic elements to produce a herarchical subdivision of the classes. From Anderson's legend, several others were obtained, adapted to the necessity and to the climatic zones at different longitudes and latitudes.

In the following, the *Global Land Cover Network* (GLCN) and the CORINE Land Cover (CLC) are presented. GLCN derives from the innovative and dynamic methodologies developed in the Land Cover Classification System (LCCS).

10.1 Global Networks in Land Cover

The main purpose of *Global Land Cover Network* (GLCN) is increasing the availability of usable and standardized information about land cover with the possibility of interchange at a global level. This kind of information is required by the institutions working in land planning and management to face urgent topics like natural and man-induced disaster mitigation, environmental protection, sustainable development, food security. The project is a common initiative between the *Joint Research Centre* (JRC) of the European Union, FAO's *Africover* Project, the *Global Terrestrial Observing System* (GTOS) and other international organizations among which is the *United Nations Environmental Programme* (UNEP).

M.A. Gomarasca, *Basics of Geomatics*, DOI 10.1007/978-1-4020-9014-1_10,

GLCN aims to provide an international, neutral and harmonized methodology for land cover classification and dynamical monitoring with wide dissemination of the produced information.

In detail the goals are the following:

- harmonization of land cover definitions and classification systems, mapping and monitoring the specifications for the maps produced with the *Land Cover Classification System* (LCCS) developed by FAO and used in *Africover* project;
- development of a standardized methodology based on LCCS;
- realization of a global land cover database organized by metadata;
- dissemination of the methodologies adopted for the land cover classification, mapping and monitoring activities and for the land planning and management-derived applications;
- information at regional and global level.

GLCN is based on the experience matured during the development of the Africover project with the definition of the LCCS.

LCCS project (Di Gregorio and Jansen, 2000) is based on this simple concept: instead of identifying pre-defined land cover classes, it is useful to select a set of universally valid classification criteria able to identify the classes. This presupposes that each class, autonomously from its typology and geographical location, can be defined by a group of pre-selected independent diagnostic attributes, the *classifiers*. The number of used classifiers determines the classification's detail level.

10.1.1 Terminology: Land Cover and Land Use

In order to produce thematic maps exploiting remote sensing techniques, it is necessary to clarify some terms often used without univocal meaning.

First of all it is important to distinguish between *land cover* and *land use*.

These two terms generate ambiguity and confusion making classifications imprecise and incomparable.

Land cover generally refers to the physical surface of the Earth, including various combinations of natural and cultivated vegetation and man-made infrastructures. Water, glaciers, rocks and bare soil and surfaces without vegetation, though being part of the terrestrial surface and not of land cover, are often considered land cover for practical reasons.

Land use, instead, includes both the way in which the Earth's biophysical attributes are modified and the reasons for which they are altered. Land use's dynamics are indicators of the land cover changes. The land use is affected by human action, especially with reference to those who decide about land management, institutions included.

Lambin et al. (1999) define the process through which the cover is modified or transformed, including two main factors:

- *activities*, or *operations*, carried out on a surface, inducing land cover transformation;
- *objectives and intentions* supporting this transformation, included the advantages in terms of products and services, and the factors inducing the use of the land, the time and the place.

Young (1994) had previously defined land use as the *management of the terrestrial surface to satisfy the human needs*, or better, *human activities that are directly correlated to the land*.

A fact apparently obvious is that different land uses can involve the same area. Nevertheless there is not a clear distinction between *cover* and *use* generating a univocal and unquestionable separation.

10.1.1.1 The Problem of the Legend

Besides land cover and land use, it is important to define the meaning of *land cover* and *land use classes*, *land cover classification* and *land cover maps' legends*. The origins of the vegetation systematic classification concept go back to the Swedish botanist Carolus Linnaeus who defined, in the mid-18th century, a classification system for plants.

Land cover classification has always been unsystematic since the first maps were produced. Maps address specific, often limited purposes, with no definition of a universal classification system. As regards this topic, it is deferred to the examination of assessments and results of several international meetings: *Meeting in the Middle* (European Commission, CCR, 2000) and *Strategies for Global Land Cover Mapping and Monitoring* (FAO/UNEP/IAO/ USAID, 2002).

The processes which led to a methodology for the definition of a Global Land Cover Network (GLCN) for classification using remote sensing systems are based on accumulated experience since Earth observation programmes started.

The following paragraphs examine the different classification systems used since 1972, beginning with the analysis of the multispectral satellite digital images era. These experiences can be generally divided into two groups:

- land cover classification systems based on pre-defined classes;
- land cover classification systems based on independent diagnostic criteria.

10.1.2 Land Cover Classification Systems Based on Pre-defined Classes and Legends

10.1.2.1 USGS Land Use/Land Cover Classification Systems (1972/1976)

This legend, the *Land Cover and Land Use Classification System*, developed by Anderson et al. (1972) at the *United States Geological Survey* (USGS), has been widely used producing land cover maps from satellite images. It is an a priori

hierarchical classification system with pre-defined land cover classes for the first two levels and, if further data are available, with the possibility to add two more detail levels (Box 10.1).

Box 10.1 Land cover/land use classification system based on remote sensing data defined by United States Geological Service (USGS) in 1976

Classification level I (mapping scale ≤ 1:250,000)	Classification level II (mapping scale 1:100,000)
1 UrbanMrt or build-up land	1.1 Residential 1.2 Commercial and services 1.3 Industrial 1.4 Transportation, communication and utilities 1.5 Industrial e commercial complexes 1.6 Mixed urban or build-up land 1.7 Other urban or build-up land
2 Agricultural land	2.1 Cropland and pasture 2.2 Orchards, groves, vineyards, nurseries, etc. 2.3 Confined feeding operations 2.4 Other agricultural land
3 Rangeland	3.1 Herbaceous rangeland 3.2 Shrub and brush rangeland 3.3 Mixed rangeland
4 Forest land	4.1 Deciduous forest land 4.1 Evergreen forest land 4.3 Mixed forest land
5 Water	5.1 Streams and canals 5.2 Lakes 5.3 Reservoirs 5.4 Bays and estuaries
6 Wetland	6.1 Forested wetland 6.2 Non-forested wetland
7 Barren land	7.1 Dray salt flats 7.2 Beaches 7.3 Sandy areas other than beaches 7.4 Bare exposed rock 7.5 Strip mines, quarries, and gravel pits 7.6 Transitional areas 7.7 Mixed barren land
8 Tundra	8.1 Shrub and brush tundra 8.2 Herbaceous tundra 8.3 Bare ground tundra 8.4 Wet tundra 8.5 Mixed tundra
9 Perennial snow or ice	9.1 Perennial snowfields 9.2 Glaciers

The criteria characterizing this classification system are the following:

- classification accuracy (see Chapter 8) minimum level 85%;
- classification accuracy uniformity for the different categories;
- possibility to repeat the method by other operators and in different dates;
- applicability of the method to very wide areas;
- the method's compatibility with multi-temporal classifications;
- definition of land cover classes to be recognized during both image interpretation and ground surveys;
- land cover classes aggregation flexibility;
- multi-temporal classifications' comparability even in the case of different land use in different periods.

USGS 1976's first level consists in nine land cover classes recognizable at 1:250,000 scale. The second level is 37 classes at 1:100,000 scale.

10.1.2.2 EarthSat GeoCover Land Cover Legend (1990)

The US Earth Satellite Corporation (EarthSat) used the USGS 1976 classification system as a base for the development of the land cover legend for the realization of the *GlobalGeoCover LC* (land cover) *database*. The legend consists of 12 classes plus the not classified, obtained by Landsat TM images digital processing for 3 years around the 1990s (Box 10.2).

Box 10.2 EarthSat GeoCover land cover legend obtained by modification of the USGS 1976

EarthSat GeoCover LC
Global Land Cover Legend

	Land Cover Class Name	*Land Cover Class Definition*
1	Deciduous forest	Trees >3 m height, canopy closure >35%, mixed con specie <25% intermixture with evergreen species
2	Evergreen forest	As above. Includes both broadleaf and needle-leaf species
3	Shrub/scrub	Woody vegetation <3 m, with at least 10% ground cover
4	Pasture	Upland herbaceous grasses 10% ground cover
5	Barren	<10% ground cover by other LC classes
6	Urban/build-up	Includes residential, commercial, industrial, transportation, sport facilities and recreation

7	Agricultural land, general	Cultivated and pasture land, except paddy agriculture
8	Rice/paddy fields	Irrigated or rain-fed
9	Wetland, herbaceous, permanent	Water table near the surface for most of the growing season. Includes playas and salt flats
10	Wetlands mangroves	Sheltered coastal (estuarine) tropical wetlands supporting woody species of mangroves
11	Water bodies	Permanent open water bodies
12	Permanent ice or snow	Includes glaciers and permanent snow fields on mountains
13	Clouds/cloud shadows/no data	Areas where land cover interpretation was not possible

10.1.2.3 National Land Cover Data (NLCD) Classification System (1992)

An example of modification of USGS 1976 system is the *National Land Cover Characterisation Project* (LCCP). The project led by USGS had the aim to develop a univocal classification in the USA defined *National Land Cover Data* (NLCD 1992); the data analysis was completed in 2000 (Vogelmann et al. 2001).

The NLCD 1992 classification scheme is not systematic-hierarchic. It is based on the combination of the modified USGS 1976 and the NOAA classification protocol called *Coastal Change Analysis Program* (C-CAP).

The USGS 1976 original number of 37 second-level classes is reduced to 21 classes (Box 10.3).

Box 10.3 National Land Cover Data (NLCD) classification system, 1992

Land Cover Classification Scheme NLCD 1992	
Digital Code	*Descriptive Name*
11	Open water
12	Perennial ice, snow
21	Low intensity residential
22	High intensity residential
23	Commercila, industrial, transportation
31	Bare/transitional
32	Quarries, trip mines, gravel pits
33	Bare rock, sand
41	Deciduous forest
42	Evergreen forest
43	Mixed forest
51	Shrub-land
61	Orchard, vineyards, other

71	Grassland, herbaceous
81	Pasture, hay
82	Row crops
83	Small grain
84	Fallow
85	Urban, recreational grasses (parks, golf courses, cemetery, etc.)
91	Woody wetland
92	Herbaceous wetland

10.1.2.4 UNEP/FAO Land Cover Legend (1993)

USGS 1976 approach also conditioned the arrangement of the classification suggested during the meeting of UNEP/FAO experts *Harmonising Land Cover and Land Use Classifications* held in Geneva in November 1993. The legend proposed (Box 10.4) consists of 9 land cover classes in the first level and 26 in the second.

Box 10.4 UNEP/FAO land cover legend, 1993

Classification level I (mapping scale $\leq$ 1:250,000)	Classification level II (mapping scale 1:100,000)
1 Inland water	1.1 River
	1.2 Freshwater lakes
	1.3 Reservoirs
	1.4 Lagoons (brackish water)
2 Swamp	2.1 Swamp
3 Barren land	3.1 Rocks
	3.2 Sand
	3.3 Ice
4 Woodland and forest	4.1 Forest
	4.2 Woodland
	4.3 Wooded grassland
	4.4 Forest plantations
5 Shrub formation	5.1 Thicket
	5.2 Bush-land
6 Heath-land	6.1 Heath-land
7 Pasture	7.1 Unimproved grassland (rangeland)
	7.2 Improved grassland
8 Cropland	8.1 Permanent crops
	8.2 Temporary crop
	8.3 Wetland crops
	8.4 Covered agricultural land
9 Build-up land	9.1 Mines and quarries
	9.2 Residential
	9.3 Commercial
	9.4 Industrial
	9.5 Transport and infrastructure

10.1.2.5 USGS and IGBP-DIS Land Cover Obtained by Modification of USGS 1976 Classification (1996)

The modified second-level USGS 1976 classification system was used for the realization of the *Global Land Cover Characteristic* (GLCC) *Database*, produced by the USGS, the University of Nebraska-Lincoln (UNL) and the JRC-SAI (*Joint Research Centre – Space Application Institute, Ispra. Italy*) within the domain of NASA programme *Earth Observing System Pathfinder*. The system was adopted by the *International Geosphere–Biosphere Programme – Data and Information System* (IGBP-DIS) in 1996 (Box 10.5).

Box 10.5 USGS and IGBP-DIS

USGS Land Cover Legend		IGBP-DIS Land Cover Legend	
1	Urban or build-up land	1	Evergreen needle-leaf forest
2	Dry-land cropland and pasture	2	Evergreen broadleaf forest
3	Irrigated cropland and pasture	3	Deciduous needle-leaf forest
4	Mixed dry-land/ irrigated cropland and pasture	4	Deciduous broadleaf forest
5	Cropland/grassland mosaic	5	Mixed forest
6	Cropland/woodland mosaic	6	Closet shrub-land
7	Grassland	7	Open shrub-land
8	Shrub-land	8	Woody Savannah
9	Mixed shrub-land/grassland	9	Savannah
10	Savannah	10	Grassland
11	Deciduous broadleaf forest	11	Permanent wetlands
12	Deciduous needle-leaf forest	12	Cropland
13	Evergreen broadleaf forest	13	Urban or build-up land
14	Evergreen needle-leaf forest	14	Cropland/natural vegetation mosaic
15	Mixed forest	15	Snow and ice
16	Water bodies	16	Barren or sparsely vegetated land
17	Herbaceous wetlands	17	Water bodies
18	Wooded wetlands	18	Interrupted areas
19	Barren or sparsely vegetated land	19	Missing data
20	Herbaceous tundra		
21	Wooded tundra		
22	Mixed tundra		
23	Bare ground tundra		
24	Snow or ice		
25	Interrupted areas		
26	Missing data		

10.1.2.6 South African Standard Land Cover Classification System (1996)

The *South African Standard* Land Cover classification system was proposed for the realization of the *National Land Cover Database* project. The first classification level consists of 10 wide land cover classes and is used as a legend for a database of aggregated land cover categories with about 400 ha ground resolution. The second classification level consists of 31 classes and was adopted to produce maps to 1:250,000 with 25 ha minimum mapping unit (Box 10.6).

Box 10.6 South African Standard Land Cover classification system

Aggregated Land Cover Classes	Full-resolution Land Cover Classes
1 Woodland and forest	1 Woodland and forest (woodland and wooded grassland)
	2 Forest (indigenous)
2 Thicket and bush-land	3 Thicket and bush-land, bush clumps, high fynbos
	4 Shrub-land and low fynbos
3 Grassland	5 Herb-land
	6 Unimproved grassland
	7 Improved grassland
4 Forest plantations	8 Forest plantations
5 Water bodies	9 Water bodies
6 Wetlands	10 Wetlands
7 Bare or degraded land	11 Bare rock and soil (natural)
	12 Bare rock soil (erosion surface)
	13 Degraded land (woodland and forest)
	14 Degraded land (Thicket and bush-land, bush clumps, high fynbos)
	15 Degraded land (unimproved grassland)
	16 Degraded land (shrub-land and low fynbos)
	17 Degraded land (herb-land)
8 Cultivated land	18 Cultivated land (permanent crops, commercial, irrigated)
	19 Cultivated land (permanent crops, commercial, dry-land)
	20 Cultivated land (permanent crops, commercial, sugar cane)
	21 Cultivated land (temporary crops, commercial, irrigated)
	22 Cultivated land (temporary crops, commercial, dry-land)
	23 Cultivated land (temporary crops, subsistence, dry-land)
9 Urban/build-up land	24 Urban/build-up (residential)
	25 Urban/build-up land (residential, thicket, forest, woodland)
	26 Urban/build-up land (residential, smallholdings, thicket, bush-land)
	27 Urban/build-up land (residential, smallholdings, shrub-land, low fynbos)
	28 Urban/build-up land (residential, smallholdings, grassland)
	29 Urban/build-up land (commercial)
	30 Urban/build-up land (industrial, transportation)
10 Mines and queries	Mines and quarries

10.1.2.7 GOFC/GOLD Land and Forest Cover Classification (1998)

Global Observation of Forest Cover (GOFC) and *Global Observation of Land Cover Dynamics* (GOLD) promoted by *Committee Earth Observation Satellites* (CEOS) have become part of the *Global Terrestrial Observing System* (GTOS).

This method is a mixed approach that combines both the land cover classification system based on pre-defined classes and legends and the diagnostic independent criteria.

These products are based on the following criteria (Box 10.7):

- classes have to be relevant for carbon study;
- land cover is based on real presence, not on potential presence;
- compatibility with already in use legends about forests;
- a hierarchical classification system compatible with multi-scale classifications;
- a flexible system able to classify different forest covers in relation to different biomass level;
- cartographic units based on forest physiognomic characteristics; floristic elements are less important than the information about the ecosystem, useful to understand the community's composition;
- compatibility between low- and high-resolution GOFC classifications.

Box 10.7 GOFC/GOLD Land and Forest Cover Classification System

Classification Level I		Classification Level II (only for the forest and flooded forest classes)			
		Level II classification criteria			
		Physiognomic		*Phenological*	*Floristic*
		Canopy cover	Tree high (m)		
1	Forest	60–100%	>2	Evergreen	Needle leaved
		40–60%	1–2	Deciduous	Broadleaved
2	Flooded forest	25–40%	<1	Mixed	Mixed
		10–25%			
3	Burned forest				
4	Water				
5	Snow an ice				
6	Barren and sparsely vegetated				
7	Build-up				
8	Cropland				
9	Grassland				

10.1.2.8 Federal Geographic Data Committee (FGDC) Standard (1998)

As with the GOFC/GOLD classification, this method is a mixed approach. The US FGDC in 1997 adopted a new reference for the classification called *National Vegetation Classification Standard* (NVCS) based on the *United States National Vegetation Classification* (USNVC) system, developed by *The Nature Conservancy*

(TNC). The purpose is to provide a systematic classification of ecological systems in the United States (Grossman et al., 1998).

NVCS is based on the combination of physiognomic–floristic analysis which allows identification of vegetation classes both top-down (divisive) and bottom-up (agglomerative).

Among the seven defined levels, the first five are based on physiognomic observations derivable from remotely sensed images, while the last two about floristic classification are based on field observations (Box 10.8).

Box 10.8 Classification levels defined by the National Vegetation Classification Standard (NVCS)

Classification level	Definition of classification level
1	*Physiognomic class*: defined by life form and vegetation cover. The first level consists of eight classes: – Closed-tree canopy – Open-tree canopy – Shrub-land – Dwarf shrub-land – Herbaceous vegetation – Non-vascular vegetation – Sparse vegetation – Non-vegetated
2	*Physiognomic subclass*: defined by – predominant leaf phenology of woody plants (deciduous, evergreen, mixed) – leaf type and periodicity of herbaceous plants
3	*Physiognomic group*: defined by a combination of factors relating to: – climate (boreal, temperate, tropical forest) – leaf morphology (broad-leaved, needle leaved)
4	*Physiognomic subgroup*: separates the natural/semi-natural vegetation from planted/cultivated types
5	*Formation level*: categorizes vegetation according the environmental/geographic factors (e.g. coastal rainforest, sub-desert evergreen shrub-land, alpine meadows)
6	*Vegetation alliances*: are characterized by one or a group of diagnostic/dominant plant species in the uppermost stratum of vegetation
7	*Vegetation association*: characterized by diagnostic/dominant plant species

10.1.2.9 CORINE Programme

The preliminary studies of the CORINE project showed that in the European Union the available information about land cover at national level was heterogeneous, fragmented and difficult to find.

Thus for correct management of environment and natural resources, a rationalization of information about land cover and its changes developing a common and trans-boundary database was considered essential.

The CORINE Land Cover Project covered large part of European and Mediterranean countries with 1:100,000 scale land cover cartography.

The origin of the project goes back to June 27, 1985 when, as suggested by the European Union (EU), the Council approved the CORINE programme, an acronym of *Coordination of Information on Environment*, experimental project for the collection, coordination and realization of information on the state of environment and natural resources within the Community (GU L 176 of 6.7.1985).

The main goals of the CORINE programme are

- to acquire an appropriate knowledge of the environment in order to address the community politics, assessing their effectiveness acts, and to integrate the environmental and political aspects;
- to unify the heterogeneous thematic cartographies of Europe produced at different levels (international, national, regional, local);
- to update the data at regular intervals: every 5–10 years.

Among the projects constituting the whole CORINE programme (biotopes, atmospheric emissions, natural vegetation, coastal erosion, water resources, soil erosion risk, etc.), the *Land Cover* addressed the research issue about the production of land cover typology maps.

The first aim of the project is the creation of a homogeneous vector database about the land cover obtained by image interpretation, classified according to a unitary 44 class nomenclature, valid for all the 27 countries of the Union, divided into three hierarchical levels (Box 10.9).

Each country can add further hierarchical levels, for example a fourth and fifth level according to specific conditions and priorities, while the first three are identical and compulsory for every one.

Based on the directions of the European Council and following up the creation of the EIONET (*European Environment Information and Observation Network*), CORINE database implementation is under the responsibility of the *European Agency for the Environment* (EEA). CORINE/LandCover90 (CLC90) and its later updates are recognized at European level by European Commission services such as DG-Regional policy, DG-Environment and DG-Agriculture and the EEA as basic instruments for the definition of political programmes related to the territory.

CORINE/LandCover90

The recommended methodology is described by the technical guide (EC-CORINE, 1994) drawn up by European Commission's experts; it includes LANDSAT 5 TM

Box 10. 9 The first three levels of the CORINE (EC) land cover classification system's legend

Level I		Level 2		Level 3	
		1.1	Urban fabric (structure)	1.1.1	Continuous urban fabric
1	Artificial surfaces			1.1.2	Discontinuous urban fabric
		1.2	Industrial, commercial and transport units	1.2.1	Industrial or commercial units
				1.2.2	Road and rail networks and associated land
				1.2.3	Port areas
		1.2.4	Airports		
		1.3	Mine, dump and construction sites	1.3.1	Mineral extraction sites
				1.3.2	Dump sites
		1.3.3	Construction sites		
		1.4	Artificial non-agricultural vegetated areas	1.4.1	Green urban areas
				1.4.2	Sport and leisure facilities
2	Agricultural areas	2.1	Arable land	2.1.1	Non-irrigated arable land
				2.1.2	Permanently irrigated areas
		2.1.3	Rice fields		
		2.2	Permanent crops	2.2.1	Vineyards
				2.2.2	Fruit trees and berry plantation
		2.2.3	Olive groves		
		2.3	Pastures	2.3.1	Pastures
		2.4	Heterogeneous agricultural areas	2.4.1	Annual crops associated with permanent crops
				2.4.2	Complex cultivation patterns
				2.4.3	Land principally occupied by agriculture, with significant areas of natural vegetation
				2.4.4	Agro-forestry areas

Level I		Level 2		Level 3	
3	Forests and semi-natural areas	3.1	Forests	3.1.1	Broad-leaved forest
				3.1.2	Coniferous forest
		3.1.3	Mixed forest		
		3.2	Shrub and/or herbaceous	3.2.1	Natural grasslands
			vegetation	3.2.2	Moors and heath-land
			associations	3.2.3	Sclerophyllous vegetation
		3.2.4	Transitional woodland shrub		
		3.3	Open spaces with little or no	3.3.1	Beaches, dunes, sands
			vegetation	3.3.2	Bare rocks
				3.3.3	Sparsely vegetated areas
				3.3.4	Burnt areas
3.3.5	Glaciers and perpetual snow				
4	Wetlands	4.1	Inland wetlands	4.1.1	Inland marshes
				4.1.2	Peat bogs
		4.2	Coastal wetlands	4.2.1	Salt marshes
				4.2.2	Salines
4.2.3	Intertidal flats				
5	Water bodies	5.1	Inland waters	5.1.1	Water courses
				5.1.2	Reservoirs
		5.2	Marine waters	5.2.1	Coastal lagoons
				5.2.2	Estuaries
5.2.3	Sea and ocean				

images and/or SPOT XS sensors interpretation. These images were radiometrically corrected and geo-referred before being distributed to the photo-interpreters (Fig. 10.1). Other sources of data were used, when available, as support to remotely sensed image analysis, such as different scale thematic maps, statistical data, different scale aero-photogrammetric images (black and white, colour and colour infrared) and other types of geographic data. The interpretation could be video-interpretation of digital images, or using printed images at 1:100,000 scale.

Additional integrations have been provided for boundary regions in order to give continuity to the classes; these products have been subsequently supervised by the European Commission's experts.

CORINE Land Cover maps production is organized in five steps (Figs. 10.1 and 10.2):

- preliminary studies;
- false-colour images production;
- computer-aided photo-interpretation;
- maps digitization;
- validation.

As regards the choice of the satellite image acquisition dates, the most recent data were acquired possibly obtained in the same year to cover the same study area and taking into account the most suitable period according to the specific vegetation phenology.

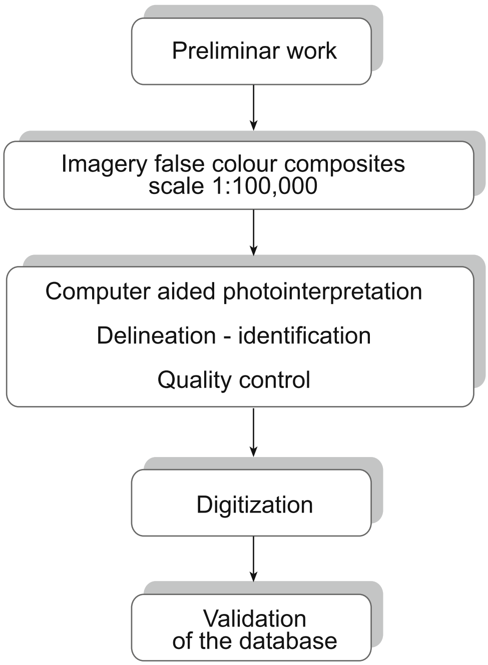


Fig. 10.1 The CORINE programme method for the production of the land cover map is organized in five phases

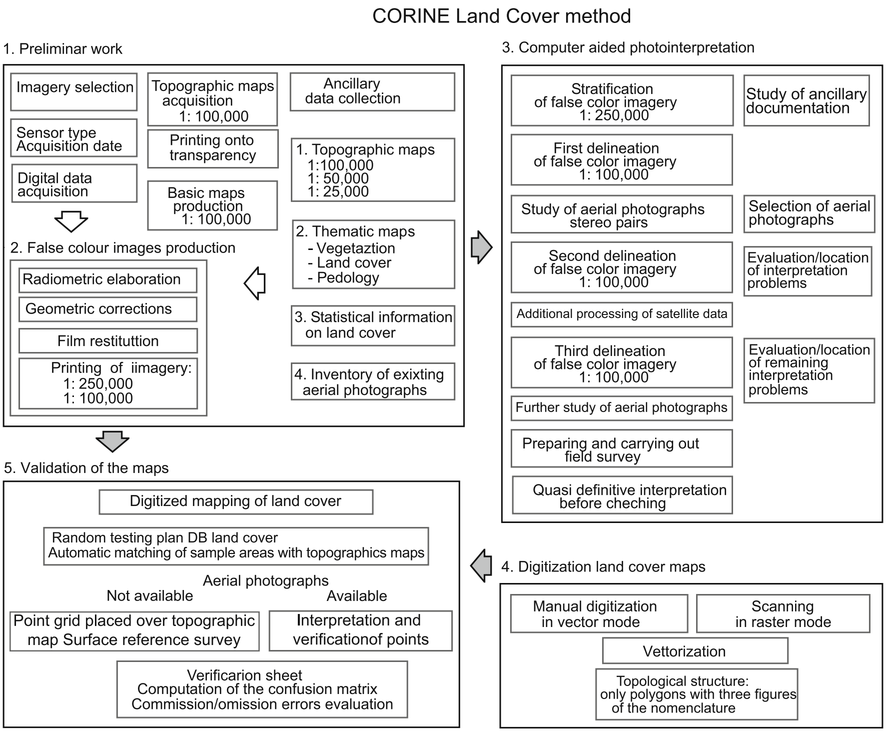


Fig. 10.2 The five phases of the CORINE/Land Cover (CLC) programme in detail

Box 10.10 CLC50 nomenclature of 4th level and some 5th level (Version 1.42, Feranec, 1995)

1 Artificial surface

1111	Areas of urban centres	1234	Sport and recreation ports
1112	Areas of ancient cores	1241	Airports with artificial surfaces of runways
1121	Discontinuous built-up areas with multiflat houses prevailingly without gardens	1242	Airports with grass surfaces of runways
1122	Discontinuous built-up areas with family houses with gardens	1311	Open cast mines
1123	Discontinuous built-up areas with greenery	1312	Quarries
12111	Industrial and commercial units	1321	Solid waste dump sites
12112	Agro-industry	1322	Liquid waste dump sites
12113	Education and health facilities	1331	Construction sites
1212	Areas of special installations	1411	Parks
1221	Road network and associated land	1412	Cemeteries

1222	Rail network and associated land	1421	Sport facilities
1232	River and lake ports	1422	Leisure areas
1233	Shipyards	1423	Recreation settlements
2	**Agriculture areas**		
2111	Arable land with large fields	2311	Intensive pastures, degraded grassland without trees and shrubs
2112	Arable land with small fields	2312	Intensive pastures, degraded grassland with trees and shrubs
2113	Greenhouses	2421	Complex cultivation patterns without scattered houses
2121	Permanently irrigated arable land	24221	Complex cultivation patterns with scattered houses
2131	Rice fields	24222	Farmsteads
22111	Vineyards with large fields	2431	Agricultural areas with significant share of natural vegetation and with prevalence of arable land
22112	Vineyards with small fields	2432	Agricultural areas with significant share of natural vegetation and with prevalence of grasslands
2221	Orchards	2433	Agricultural areas with significant share of natural vegetation and with prevalence of scattered natural vegetation
2222	Berry fruit plantations	2434	Agricultural areas with significant share of ponds and with prevalence of scattered natural vegetation
2223	Hop plantations	2435	Agricultural areas with significant share of permanent crops and with the presence of scattered natural vegetation
2226	Wild willow plantations		
3	**Forest and semi-natural areas**		
3115	Plantations of broad-leaved forests	3244	Forest nurseries
3121	Coniferous forests with continuous canopy	3245	Damaged forests
3125	Plantations of coniferous forests	3312	Dunes
3131	Mixed forests created by alternation of single trees with continuous canopy	3313	River banks
3135	Mixed forests created by alternation of stands of trees with continuous canopy	3321	Bare rocks
3139	Plantations of mixed forests	3331	Sparse vegetation on sands or loess
3211	Natural grassland prevailingly without trees and shrubs	3332	Sparse vegetation on rocks

3212	Natural grassland with trees and shrubs	3333	Sparse vegetation on salines
3241	Young stands and clear-cuts	3341	Burnt areas
3243	Bushy woodlands, natural regeneration areas		
4	**Wetlands**	**5**	**Waters**
4111	Freshwater marshes	5111	Rivers
4113	Saline (alkali) inland marshes	5112	Channels
4121	Explored peat bogs	51211	Natural water bodies with continuous water supply
4122	Natural peat bogs with scattered trees and shrubs	51212	Natural, temporary, salt-affected water bodies
		51221	Artificial lakes, reservoirs
		51222	Fish ponds

In order to support the interpretation phase, in CLC90 a good choice of spectral data was offered by the availability of Landsat (MSS and TM) and SPOT (HRV) as well as multi-temporal acquisitions. Ancillary material collection and preparation and the realization of maps supporting the video-photo-interpretation were also initial steps of the process.

False-colour image interpretation was preceded by some steps aimed at reducing the errors in the acquisition phase. The *system errors'* correction concerned the remove of strip effects (de-striping), the correction of radiometric anomalies caused by saturated pixels (black or white strips) or salt and pepper effect adopting suitable filtering procedures. Moreover, due to the necessity to overlap the images with other thematic layers, the data were georeferenced and geometrically corrected.

False-colour composites, using the most significant spectral bands, were mostly used as highlighted in Table 10.1, due to their importance in vegetation discrimination, working both on video and on photographic prints.

Table 10.1 Multispectral colour composites of some satellite images mostly used for land cover interpretation

	Electromagnetic spectral band		
Sensor	Red	Green	Blue
Landsat MSS	Band 7 (0.80 – 1.1 μm)	Band 5 (0.60 – 0.70 μm)	Band 4 (0.50 – 0.60 μm)
Landsat TM	Band 4 (0.76 – 0.90 μm)	Band 5 (1.55 – 1.75 μm)	Band 3 (0.63 – 0.69 μm)
Landsat TM	Band 4 (0.76 – 0.90 μm)	Band 3 (0.63 – 0.69 μm)	Band 2 (0.52 – 0.60 μm)
Spot XS	Band 3 (0.79 – 0.89 μm)	Band 2 (0.61 – 0.68 μm)	Band 1 (0.50 – 0.59 μm)

Box 10.11 reports one of the interpretation keys used in RGB:432 colour composition of Landsat TM images.

Box 10.11 Interpretation sheet based on the false-colour composite of Landsat ETM+ RGB:432

Land cover typology	False-colour composite (RGB 432)
Urban fabric (structure)	Blue (depending on the density from dark to light)
Dump site, bare rock, sand, dune, concrete infrastructure	White
Networks (road, railway)	Dark blue
Permanent ice or snow (clouds)	White, white-bluish
Annual crops	Red (growing phase), grey–pink (after harvesting), blue–white (ploughed field)
Permanent crops	Red–pink
Deciduous forest	Bright red
Coniferous forest	Brown–red
Grassland	Bright pink, bright red
Wetland	Black or dark red
Rangeland, shrub	Grey–yellow, grey–pink, light brown
Burnt areas	Black, dark grey, bluish

Photo-interpretation was carried out on satellite imagery overlapping thematic layers as hydrography. The operation was realized in an interactive way, being the ancillary images and the topographic base available at any time, as well as video false-colour composites and their possible processing: filters, contrast enhancement, vegetation and soil indices, supervised and unsupervised classifications, statistical analyses, principal components analysis, etc.

The procedure included the following steps:

- tracing the boundary of each area recognized as belonging to a single cover (*delineation*);
- assignation to each area with a code corresponding to a class, aided by support data (*identification*);
- extrapolation of the two previous operations on the zones of the images presenting similar characteristics: colour, structure and texture (*generalization*).

This non-linear procedure implied interactive controls and constant comparisons with ancillary information, as shown in Plate 10.1 where the case of Marseille is reported.

Once traced, the boundaries of each class had to be digitized in order to be inserted in a database including the thematic layers of the CORINE project's natural resources (Fig. 10.3; Table 10.2)

The procedures used for the digitization were

- *manual*, by proper software, carried out by assigning each polygon with a code corresponding to a cover class;
- *automatic scanning*, with data restitution in raster format, followed by a process of map harmonization.

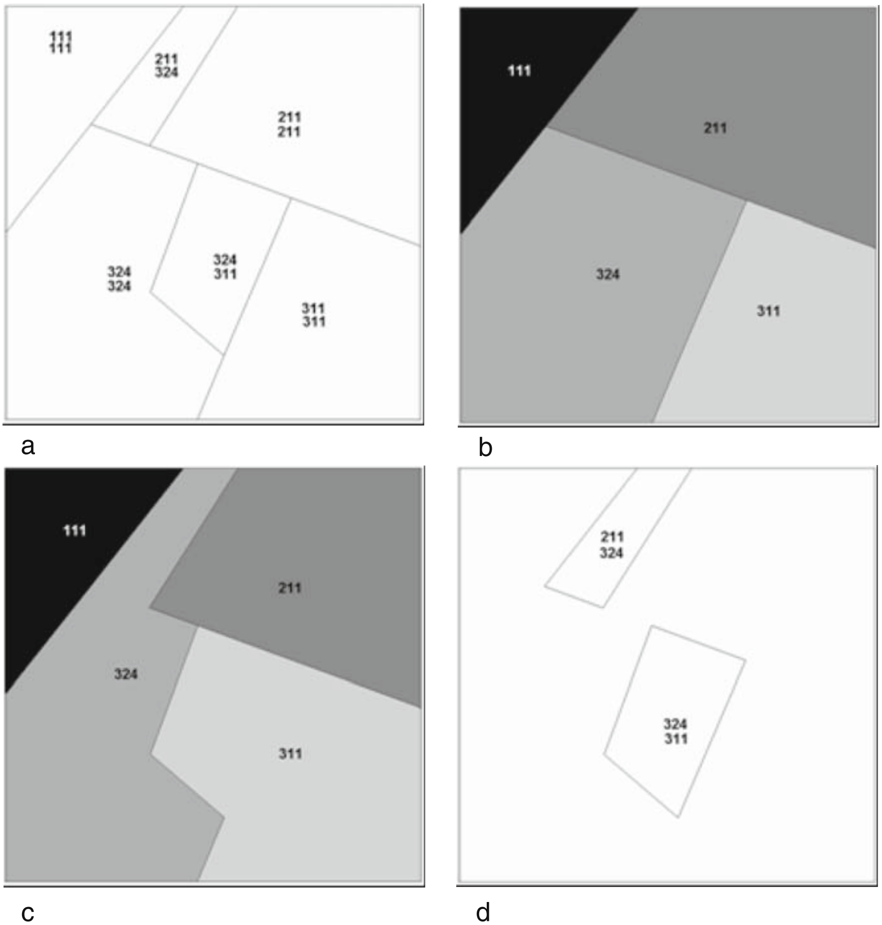


Fig. 10.3 CLC database structure: (**a**) structure of the imagery used in qualitative interpretation, with polygons code in 1990 (above) and in 2000 (below); (**b**) derivation of the correct version of CLC90; (**c**) derivation of the correct version of CLC2000; (**d**) CLC-change

CORINE/LandCover2000

The CORINE Land Cover update to the year 2000 (CLC2000), called *Image and Corine Land Cover 2000* project (I&CLC2000), schedules the acquisition of remotely sensed Landsat 7 ETM+ images (*Image 2000*), the production of the geographical database of land cover for the year 2000 and the assessment of the land cover/land use changes between the years 1990 and 2000 (*CLC-change*).

In order to realize these two products, the revision of CLC90 imagery was necessary together with the acquisition of new historical Landsat 5 TM images referred to the year 1990. Essential pre-requirement for the evaluation of land use changes is in fact the geometric and thematic consistency between the CLC90 and the CLC2000.

Table 10.2 Main characteristic of the projects CORINE Land Cover 1990 (CLC1990) and of the most recent Image and CORINE Land cover 2000 (I&CLC2000)

Characteristics	CLC1990	I&CLC2000
Satellite imagery temporal range (1° draw)	1986–1995	–
Satellite imagery temporal range (updating)	1988–1992 (Image90)	1999–2001 (image2000)
Cartographic georeferencing system		UTM32N datum WGS84
Map scale	1:100,000	1:100,000
Minimum unit CLC	25 ha	25 ha
Minimum unit CLC-change	–	5 ha
Geometric accuracy (RMSE)	50 m	25 m
Satellite imagery	100 m	<100 m
CLC		
Metadata	Incomplete Non-standard	Standard
Policy of data distribution	Not defined	Defined

CORINE/LandCover2006

The implementation of CLC2006 focuses mainly on the identification and mapping of land cover changes between 2000 and 2006. CLC2006 is scheduled as a direct continuation of previous Corine land cover mapping campaigns. Land cover changes larger than 5 ha must be mapped, regardless of location. As in previous years, the work is a shared approach between EEA and EU member countries. While the image acquisition and pre-processing is centrally organized, the land cover mapping is done by member states to benefit from local knowledge.

CORINE/LandCover: The Case of Italy

The Corine/Land Cover map has been produced for the 27 countries of the European Union. As an example, the case of Italy is presented to clarify the general approach used in Europe.

The original interpretation of CORINE/Land Cover for the year 1990 (CLC90) did not have a homogeneous implementation in Italy nor a standardized methodology of realization; this product has been realized by several authorities. Moreover, several problems were faced in the interpretation process because, as in the other European countries, the data set used for CLC90 was constituted by images acquired over a large interval of time (1980–1992).

In order to overcome some of these problems during the project *Image and Corine Land Cover 2000* (I&CLC2000), different geographical databases, for the quantification of the processes of evolution of land use/land cover, have been realized for 1990 and 2000 in Italy.

The results of the project constitute a progress in comparison to the first layout CLC90 and a fundamental instrument for the knowledge of the environment and for decision making, planning and land management support.

The basic information used for the CLC2000 database is generated from:

- Image90: mosaics of Landsat 5 TM images from 1980 to 1992;
- Image2000: mosaics of Landsat 7 ETM+ images around the year 2000 (± 2 years);
- digital terrain model with interpolation resolution of 75 m;
- National digital ortho-photographic acquisition in natural colour, IT2000 in scale 1:10,000 of the project Terra-Italy (CGR, Parma 1998–1999) with ground geometric resolution of 1 m;
- topographic cartography IGM 1:250,000 raster;
- vector map of land use and land cover in scale 1:250,000;
- local cartography of land use/land cover in scales 1:10,000–1:50,000;
- vector land use map ISTAT 1991 in scale 1:50,000.

The use of the digital orthophotos IT2000 (1:10,000) has been important but limited, for incongruities of scale, to the thematic interpretation of doubtful cases in the year 2000, and not for the definition of the geometry of the polygons.

The I&CLC2000 project has been articulated in two interconnected phases:

- *Image2000: acquisition* of a European mosaic of images Landsat 7 ETM+ and their ortorettification, with positioning error, Root Mean Squared Error (RMSE), less than 25 m, co-registration of CLC90 with Image2000 acquiring at least 1000 ground control points for each Italian region;
- *CLC2000:* activities related to the mapping and to the interpretation of the changes of land use/land cover. A revisited and correct version of CLC90, the update of the new *CLC2000* and the production of the *CLC-change*.

Land cover products consist of a vector digital cartography 1:100,000 scale of representation. The minimal mapping unit (MMU) is 25 ha, equivalent at scale 1:100,000 to a circle of 2.8 mm or a square of 5 mm × 5 mm) and the minimum polygons width is 100 m (1 mm at 1:100,000). Finally the geometric accuracy is 100 m.

In the *CLC-change* product, the MMU is 5 ha and the scale of representation 1:50,000.

The legend system adopted for I&CLC2000 is coincident with the CLC90 nomenclature (Box 10.9).

Besides the standard products, CLC2000 and CLC-change, other products have been realized in:

- *Image90 and new CLC90*: Image 90 is satellite imagery archive representative of the year 1990 at national level. This data set, constituted by a mosaic of Landsat 5 TM images co-registered with Image2000, is the basic source for the revision of CLC90 necessary to guarantee the new qualitative standards of I&CLC2000;
- *CLC2000_IV*: CLC2000 with the introduction of the interpretation of a fourth thematic level (IV) for the forest areas and the other natural and semi-natural environments;
- Evaluation of the classification thematic accuracy of CLC2000.

Finally, a policy of distribution and access has been defined at national level for the three definitive land cover products *CLC2000, CLC-change and CLC2000_IV*.

Like all the information obtained by complex procedures, CORINE maps produced in different periods contain errors in the attribution of the classes as well as in drawing the polygons' boundaries; thus it is necessary to provide assessment of the products' reliability. The validation procedure scheme involved the comparison of the obtained database information with a new data set, coming from independent in-field surveys or derived by aerial photos not used during the photo-interpretation phase.

The assessment has been realized by a statistical approach using representative sample surfaces randomly selected. The assessment defines the percentage for each land cover category correctly classified and the overall accuracy (see Chapter 8).

10.1.3 Land Cover Classification Systems Based on Diagnostic Independent Criteria

The comparison between classification systems based on pre-defined classes highlights the limit of the method for both regional and global applications. Due to excessive rigidity, these classification systems are often limited to defined geographical areas and drawn for a specific kind of application. The systems described present a lack of clearness in biomass classes' definition, and a certain incompatibility among them, often making the relationship inconsistent. For this reason, none of the mentioned systems can represent a standard method applicable anywhere.

The only system based on independent and universal diagnostic land cover criteria is the *Land Cover Classification System* (LCCS) developed by FAO/UNEP. The description of LCCS is reported later in this chapter in the *Africover project*.

At the moment, LCCS is the most accredited to become the international standard classification system due to its flexibility, applicability in any climatic region and environmental condition and compatibility with the existing classification systems.

10.1.3.1 Africover Project

Origins and Objectives

This project was conceived to answer African countries' numerous needs of having available georeferenced geographic information about natural resources at national, regional and local scale.

The basic idea of the project is to realize a land use/land cover thematic cartography for very large areas exploiting satellite image interpretation and the use of satellite positioning systems, GPS, for field data location. The methodology must, as far as possible, be automatic in order to reduce processing times, allowing integrations with other cartographies and providing continuously up-to-date products that can be used in different domains.

A basic point of the project is the distinction, already treated at the beginning of this chapter, between:

- *land cover*: what is physically recognizable on the surface in a certain moment, and thus photo-interpretable;
- *land use*: land cover function in relation to human activities, thus including socio-economical information.

One of the objectives of the *Africover* project, conceived by the Food and Agriculture Organization (FAO) of the United Nations, is to organize a digital and geo-referenced database of the land cover with a geographic base (international and administrative boundaries, topography, hydrography, road network, toponyms) at 1:200,000 or 1:100,000 scales in function of the size of the area of interest. National databases are harmonized at regional level allowing the generalization of the information related to the required detail of the application.

A second goal is to train experts, at national and regional level, to use the geographical database at local level and to update it periodically.

The first operative module includes 11 countries in the Nile River basin: Burundi, Democratic Republic of Congo, Egypt, Eritrea, Kenya, Rwanda, Somalia, Sudan, Tanzania, Uganda and Ethiopia (see Plate 10.2). The project started in January 1997 thanks to *Italian Ministry of Foreign Affairs* governmental funds and based in Nairobi, hosted by the *Regional Centre for Services in Surveying, Mapping and Remote Sensing* (RCSSMRS), in cooperation with the *Soil Resources Management and Conservation Service* (SRMCS).

The area covered by the study is equal to 8.5 million km^2, almost one-third of Africa, and to realize it more than 400 Landsat TM 5 images have been processed and analysed.

The final users of *Africover* are technical experts and decision makers involved in natural resource management and monitoring like ministries (rural development, environment, agriculture, forestry, etc.), international development agencies (*United Nations* [UN], *World Bank* [WB] *European Union* [EU], *International Fund for Agricultural Development* [IFAD]), intergovernmental organizations (*Intergovernmental Authority on Drought and Development* [IGADD], *Southern African Development Community* [SADCC], *Comité inter-Etats de Lutte contre la Sécheresse au Sahel* [CILSS]), bilateral cooperation agencies, non-governmental organizations and private national operators.

The Components of the *Africover* Project

The compilation of a technical specification has been based on the following fundamental components:

- definition of the classification hierarchical procedure;
- definition of the rules about datum, ellipsoid, maps projection, altimetric and planimetric accuracy on the African continent;
- elaboration of project supervision and assessment methodologies.

The complex structure of image-interpretation process includes several components suitable for the definition of technical specifications and of a methodology for Africa's land cover mapping (1:200,000 or 1:100,000, generalizable to 1:1,000,000). The classification methodology includes the development of:

- cartographic concepts and standards for the MADE (*Multipurpose Africover Database for Environment resources*) database;
- land cover classification system, LCCS;
- system for the interpretation, processing and visualization of remotely sensed images: AIMs, *Africover Interpretation and Mapping System*, and GeoVIS, *Geographic Vector Interpretation System*;
- management instrument of ADG (*Africover Database Gateway*) database;
- interactive guide for photo-interpretation, AID: *Africover Interactive Database for interpretation*;
- interactive statistical software tool, MAP: *Map Accuracy Programme*, to assess maps' accuracy.

The coordination phase of the involved countries to establish and to apply the standards for information, instruments, analysis methods and procedures proved very complex. The standardization process involved the integration of national, regional and local specifications in one system.

Methodological Approach

Africover's methodology is based on the following characteristics:

- *uniform detail level* for each aspect of the land cover (agriculture, woods, rangelands, etc.);
- *neutral and univocal definition* of *thematic classes* in order to be suitable for a wide number of users belonging to different domains;
- *accessibility* to *Africover* products for qualified personnel who request them, also by distribution through the Web;
- *consistency* in terms of homogeneity with similar land cover products realized for other areas of the African continent;
- *quality* and *accuracy* uniformly distributed in Africover products, even in regions very different one from the other.

Multipurpose Africover Database for Environmental resources (MADE) is the container of basic information for the GIS users who want to access environmental information. It is constituted by a set of detailed and homogeneous information about environment and land cover which can be used by a wide number of specialized users (see Plate 10.3).

This database structure has the advantage to reduce the costs and to improve local, regional and national level efficiency in resource management.

MADE's repository in each country is a governmental institution (*National Focal Point*) designated for data distribution and customization in several domains:

- landscape planning;
- forestry;
- rangelands;
- agriculture;
- statistics;
- wildlife.

Land Cover Classification System (LCCS)

The *Land Cover Classification System* (LCCS) is based on the concept of defining a land cover class through universally valid criteria.

LCCS is an Expert System, programmed in Access in its first version and in Visual Basic in the second, which allows the user to create land cover typologies and save them in a specific database called *Legend*.

LCCS intends to conceptualize, to define and to classify the land cover (Di Gregorio and Jansen, 1996; FAO 1997).

This presupposes that any vegetation cover class, independently of its geographical position, can be defined by different pre-selected independent diagnostic attributes. The number of used classifiers determines the classification level of a cover. The level of the detail of classification determines the number of classifiers to be provided

The implemented software has the following functions:

- to classify the land cover in a systematic way and with great flexibility. The system has been developed considering the possibility of describing several land cover typologies. The concept is that the possible classes have not been created a priori, however the system allows creation of new classes if needed;
- to record the ground surveys that are automatically collected together with ancillary data;
- to correlate and compare classifiers and legends allowing links between land cover classes and classifiers, i.e. the level of independent and diagnostic parameters used to define a class.

The programme is constituted by three modules (Box 10.12):

Box 10.12 The Land Cover Classification System (LCCS) is organized in three modules inter-related to each other

Classification module	Legend module
Classifiers and attributes	Legend creation by storing, in a hierarchical structure, the land cover classes defined in the classification Module
Glossary	Class description and definition
Definition of land cover classes according with: - an initial Dichotomous Phase - a subsequent Modular-Hierarchical Phase	Land cover classes creation defined by the user
At any level the user can ask for the land cover class and store its Boolean formula, numerical code and class name in the Legend Module	Legend visualization
	Print and transfer

Translator Module			
Translation of the existing legends and classifications in the LCCS language	Assessment of similarity of classes according to other legends and classifications using LCCS	Comparison of classes of translated legends and classifications and their attributes using LCCS as a reference base	Comparison of two land cover classes of LCCS and their attributes (land cover class and in-field survey)

- *classifier*: used to create new classes in the selected legend; land cover classes are defined by the combination of independent criteria, environmental attributes and specific technical attributes hierarchically ordered;
- *legend*: list of land cover classes according to the domain to which they belong. This module exports data into a file allowing the user to add more classes and standardize them;
- *translation*: any external legend of other thematic cartographies can be translated into the LCCS.

The Classification Module

The LCCS is an a priori system, conceived to answer the needs of a wide variety of users, and it is independent from the scales and purposes for which it is used.

This system divides the classifiers into eight groups corresponding to the major land cover classes in the world. This choice strongly reduces the number of classifiers and significantly reduces the classification procedures.

This approach involves LCCS implementation in two phases:

- dichotomous classification at the beginning; followed by
- modular-hierarchical classification.

The dichotomous phase uses the following classification criteria, or levels (Box 10.13):

Box 10.13 Definition of the three levels of the FAO-LCCS dichotomous classification system

Dichotomous phase		
Initial-level distinction Presence–absence of vegetation	*Second-level distinction Edaphic condition*	*Tertiary-level distinction Artificiality of cover*
Primarily vegetated	Terrestrial	A11 Cultivated and managed terrestrial areas
		A12 Natural and semi-natural vegetation
	Aquatic or regularly flooded	A23 Cultivated aquatic or regularly flooded areas
A24 Natural and semi-natural aquatic or regularly flooded vegetation		
Primarily non-vegetated	Terrestrial	B 15 Artificial surfaces and associated areas
		B16 Bare areas
	Aquatic or regularly flooded	B27 Artificial water bodies, snow and ice
B28 Natural water bodies, snow and ice		

- presence of vegetation;
- soil mineral conditions;
- antropic modification of vegetation cover (cultivated and/or managed) (biomass artificiality) which leads to the eight cover classes of the third level.

The first phase is called *dichotomous* as the user has to choose between two options: vegetated or non-vegetated, terrestrial or aquatic vegetation, natural or cultivated vegetation. The dichotomous phase leads to eight modules including the main land cover types:

- Cultivated and managed terrestrial areas
- Natural and semi-natural terrestrial vegetation
- Cultivated aquatic or regularly flooded areas
- Natural and semi-natural aquatic or regularly flooded vegetation
- Artificial surfaces and associated areas
- Bare areas
- Artificial water bodies, snow and ice
- Natural water bodies, snow and ice

Each of these is then subsequently classified in the modular-hierarchical phase. Each of them has its own group of pre-defined classifiers, also called cover indicators, chosen in function of the class cover type. These pure classifiers have a pre-defined hierarchical structure that must be followed during the classification procedure. Some examples are reported in Boxes 10.14, 10.15, 10.16 and 10.17.

Box 10.14 Modular-hierarchical classification for the class natural and semi-natural terrestrial vegetation

Classifiers	1° hierarchical module	2° hierarchical module	3° hierarchical module
Pure classifiers of land cover	Physiognomic-structural aspects: Life Form	Woody	Trees
			Shrubs
		Herbaceous	Forbs
			Graminoids
		Lichens–Mosses	Lichens
			Mosses
	Morphological and phenological characteristics	Broadleaf	Evergreen or deciduous
		Needle-leaf	Evergreen or deciduous
		Aphyllous	Annual
			Poli-annual
Environmental attributes	Landform		
	Lithology		
	Climate		
	Altitude		
	Erosion		
	Cover and density		
Specific technical attributes, e.g. floristic aspects	Single plant species	Dominant species	
		More frequent species	
	Groups of plant species	Statistically derived plant groups	
		Plant groups derived without statistical methods	

Box 10.15 Example of pure classifiers for the class natural and semi-natural terrestrial vegetation

	Natural and semi-natural terrestrial vegetation			
Classification levels	*Classifiers (diagnostic attributes)*			
4	Principal level type: arboreous, herbaceous	Vegetation coverof the principal level	Vegetation height	Spatial distribution: texture
5	Leaf type: broadleaf, needle-leaf, aphyllous		Leaf phenology: evergreen, deciduous, mixed	
6	Vertical stratification 2nd–3rd level		Vegetation cover 2nd–3rd level	Vegetation height 2nd–3rd level

Box 10.16 The modular-hierarchical phase: example of classifiers and attributes selection for cultivated and managed terrestrial lands (left) and natural and semi-natural aquatic or regularly flooded vegetation (right). Same cap letter corresponds to same classifier selection

Cultivated and managed terrestrial lands				Natural and semi-natural aquatic or regularly flooded vegetation		
A Life form	B Spatial aspects			A Life form and cover	B Height	
C Crop combination			*Pure Land cover Classifier*	C Water seasonality		
D Cover-related cultural practices				D Leaf type Life cycle	E Leaf phenology	
					F/G/H Stratification	
L Landform	M/N Lithology	O Climate	*Environmental attributes*	L Landform	M/N Lithology	O Climate
P Altitude	Q Erosion	W Cover and density		P Altitude	Q Erosion	W Water quality
S Crop type			*Specific Technical attributes*	T Floristic aspects		

Class description can be progressively improved by increasing detail levels and can stop at any of them according to the available information. Besides physiognomic and phenologic classifiers, there are two optional categories:

- environmental attributes (land morphology, lithology, climate, height, erosion, cover density) and
- specific technical attributes (e.g. crop type, floristic aspects, etc.).

For instance, the physiognomic–structural classifiers for the class *natural and semi-natural terrestrial vegetation,* are shape and cover, height and spatial distribution, while the morphologic–phenologic classifiers are leaf typology, cycle, phenology and stratification.

The classifiers are grouped in blocks, the so-called *levels*, in which the choices are mutually exclusive.

The same procedure is repeated for each attribute until the specific technical level.

At the end of this phase, the programme updates the database generating specific codes which indicate the chosen class (Box 10.17). The code is composed of

Box 10.17 Formation of land cover classes in the phase of hierarchical modular classification. Examples of the class: natural and semi-natural terrestrial vegetation (A12)

Classifier	Boolean formula	Standard class name	Numeric code (GIS)
Type (T)	A3	Forest	20004
Vegetation cover (C)	A3A10	Closed forest	20005
Height (H)	A3A10B2	High closed forest	20006
Spatial distribution	A3A10B2C1	Continuous closed forest	20007
Leaf type	A3A10B2C1D1	Broad-leaved closed forest	20095
Leaf phenology	A3A10B2C1D1E2	Broad-leaved deciduous forest	20097
Second level (T,C,H)	A3A10B2C1D1E2 F2F5F7G2	Multi-layer broad-leaved deciduous forest	20628
Third level (T,C,H)	A3A10B2C1D1E2 F2F5F7G2 F2F5F10G2	Multi-layer broad-leaved deciduous forest with emergents	20630

- a Boolean formula, consisting in an alpha-numerical string referring to the used classifiers;
- a uniformed name of the class;
- a numerical code, used in GIS internal applications.

Legend Creation

This module shows the list of classes related to an available legend (chosen among those archived in the Access-format file) and allows elimination of those classes considered useless. Two other elements can be added to better specify the classes: the class's tag and a personalized description (Box 10.18).

Each class presents a *standard class description* that can be visualized and contains both the Boolean code and the additional information provided by the interpreter.

Moreover, the list of classifiers and their description can be visualized in the legend, in order to interpret the classes. This functionality is particularly useful when some queries are launched in a GIS grouping some classes' categories. The information contained in the legend can be modified, saved, printed, imported and exported.

A legend can be modified adding the necessary classes; specific classifiers, for each class, must be selected in the classification module. It is always possible to add new attributes following the pre-defined standards.

Translator

This module offers the possibility of:

- transferring existing classifications into LCCS language;
- evaluating similarities among classes belonging to different classifiers by using LCCS criteria;
- comparing classes of other classifiers, legends and their attributes by using LCCS as a reference;
- comparing two LCCS classes, one derived from a preliminary interpretation and the other from ground data.

The Image Interpretation and Elaboration System (GeoVIS)

GeoVIS (*Geographic Vector Interpretation System*) is the program dedicated to satellite images processing, analysis and interpretation. It replaces the original AIMs, i.e. *Africover Interpretation and Mapping System*, raster system for images digital processing, based on a vector digitization software tool with advanced functions for the video digitization of the polygons. This program supports the interpretation activities reducing, at the same time, the GIS work needed to build and maintain the database. GeoVIS second version, operative since 2003, allows high

flexibility in raster data management (image visualization in four synchronized windows), the possibility of using automatic polygons digitization instruments (magic wand: automatic classification per single class) and the possibility of integrating external data (automatic classification, already existing cartography, etc.) within the working database.

GeoVIS specific functionalities and functions are

- ergonomic and complete environment for video-interpretation;
- direct link with LCCS;
- data compression capacity;
- topological system with link to other thematic layers;
- link to the database with capacity of queries;
- easy to use.

The programme is related and completely integrated with the LCCS classification module. When the user draws a polygon in relation to a specific class in the database, a code is automatically created and attributed.

Africover Database

In concert with LCCS, *Africover Database Gateway* (ADG) module assures a wide range of functionalities in the database. The query can be formulated either on the basis of land cover polygons classification name or exploiting classifier codes (attributes and indicators).

The user can create a new classification by choosing a group of attributes in function of the final use.

The available operations in this module are the following:

- multi-user structure for a flexible data query; selecting the attributes according to individual concern in a defined area of interest (administrative, geographic);
- definition of the user name in order to organize a user-defined classes' codes;
- statistics about the areas;
- immediate visualization of the resulting areas;
- data output in different GIS vector formats.

Africover Interactive Database (AID)

Africover Interactive Database (AID) is an interactive program to guide and standardize land cover interpretation. Once different types of images, ground truth, ancillary data and other information are selected, the data are organized in an interactive way to guide the user in the identification of land cover classes.

Box 10.18 Typologies of land cover with their structural domain. Distinction at the third level of the dichotomous phase into eight major land cover types

Typology of land cover	Structural domain
A11. Cultivated and managed terrestrial areas	Trees: broad-leaved, needle leaved, evergreen, deciduous Shrubs: broad-leaved, needle leaved, evergreen, deciduous Herbaceous: graminoids, non-graminoids Arable land
A12. Natural and semi-natural vegetation	Forest Woody Closed shrubs Sparse shrubs Pasture Sparse vegetation Lichens–Mosses
A23. Cultivated aquatic or regularly flooded areas	Graminoids (rice, reeds) non-graminoids woody
A24. Natural and semi-natural aquatic or regularly flooded vegetation	Forest Woody Closed shrubs Sparse shrubs Pasture Sparse vegetation Lichens–Mosses
B15. Artificial surfaces and associated areas	Built-up − linear: roads, railways, communication lines and pipelines Non-linear: industrial, urban Non-built-up − waste dump deposits, extraction sites
B16. Bare areas	Consolidated: bare rock and/or coarse fragments, gravel, stones and boulders, hardpans, ironpan/laterite, petrocalcic, petrogypsic Unconsolidated: bare soil or other unconsolidated material, stony, loose and shifting sands
B27. Artificial water bodies, snow and ice	Reservoirs artificial snow artificial ice
B28. Natural water bodies, snow and ice	Natural basins Natural snow Natural ice

Its main characteristics are

- interactive database for photo-interpretation;
- photo-interpretation procedures' detailed description;
- interpretation of land cover characteristics;
- precise description of LCCS classes;
- link interface between in-field data, photographs, aerial photos, vegetation indices, land morphology and soil geology.

A synthetic frame of Africover classification results applying the Land Cover Classification System (LCCS) methodology is reported in Plate 10.4. Figures *a* and *b* show the aggregated land cover database for Kenya and Sudan.

10.2 Summary

This chapter introduces the legends and nomenclatures used in the recent past and conceived in modern times for land use/land cover classification. Between the numerous thematic maps that can be realized from remote sensing data, the attention is focused on land use/land cover as it is the most widely diffused and required also in relation to its importance as a basic reference for the development of other themes.

Even though the presentation of nomenclatures' different typologies and related thematic classes might appear excessive, their reading helps to understand the different philosophy adopted in the creation of the map. The legends are based on a hierarchical model, with classes changing in function of the classification purposes and the specific region in which they are mostly used; however, a new method has recently been proposed to obtain more exhaustive classifications, adaptable to any geographical situation, updatable, dynamic and able to use the thematic cartography produced in the past.

The most used nomenclature is defined by Anderson at al. (1972), from which many others have been derived such as the *CORINE Land Cover* (CLC) Legend approved by the European Commission in 1985. CORINE (*Coordination of Information on Environment*) is a hierarchical and rigid legend which does not always satisfy the user's need but which has the great advantage of uniformly representing the land cover of European Union countries. Land and natural resources' sustainable management and environmental protection require the periodic assessment of land use/land cover dynamics; thematic cartography produces a large quantity of information that is often not easy to use due to not being updated and not being integrated with other sources of data. For this reason, an international land cover classification system universally recognized, flexible and dynamic to respond to even very different needs is much needed.

Responding to these needs, FAO (*Food and Agriculture Organization of the United Nations*) and UNEP (*United Nations Development Programme*) jointly

developed the *Global Land Cover Network* (GLCN). This initiative represents a first positive experience based on harmonized and shared programmes with the aim of producing interoperable information on land cover.

Further Reading

Anderson J.R., Hardy E.E., Roach J.T., 1972, A land-use classification system for use with remote sensor data. US Geological Survey Circular 671, Washington DC, USA, p. 16.

Di Gregorio A., Jansen L.J.M., 2000, Land Cover Classification System (LCCS). Classification Concepts and User Manual. Food and Agriculture Organisation of the United Nations (FAO), Rome, Italy, pp. 80+99 of appendices.

Heymann Y., Steenmans Ch., Croissille G., Bossard M., 1994, CORINE Land Cover. Technical Guide. Office for Official Publications of the European Communities, EUR12585 EN, Luxembourg, p. 137.

Lambin E.F., Baulies X., Bockstael N., Fischer G., Krug Leemans T.R., Moran E.F., Rindfuss R.R., Sato Y., Skole D., Turner B.L. II, Vogel C. 1999, Land-Use and Land-Cover Change (LUCC). Implementation Strategy (IGBP Report 48, IHDP Report 10). IGBP and IHDP Secretariats, Stockholm, Bonn, p. 125.

Bibliography

Alberts J., Tauch R., 1994, Mapping from space - cartographic applications of satellite image data. GeoJournal, 32(l): 29–37.

Belward A.S., 1996, The IGBP-DIS Global 1-Km Land Cover Data Set (DISCover): Proposal and Implementation Plans Toulouse, France, A.S. Belward (Ed.) IGBP-DIS, Toulouse, Fran Working Paper No. 13, p. 61.

Belward A.S., Estes J.E, Kline K.D., 1999, The IGDB-DIS global 1 km land cover data set DISCover: a project overview. Photogrammetric Engineering and Remote Sensing, 65(9): 1013–1020.

Bossard M., Feranec J., Otahel J., 2000, CORINE land cover-technical guide – Addendum 2000, European Environment Agency, Technical report n. 40.

Büttner G. et al., 1998, The European CORINE Land Cover Database, ISPRS Commission VII Symposium, Budapest, September 1–4, 1998, Proceedings, pp. 633–638.

Chodota M.W.L., 1996, The FAO Africover Project and a Possibility of a Unified Geodetic Datum for Africa (UGDA). Presented at the Ninth United Nations Cartographic Conference for Africa. United Nations Economic Commission for Africa (ECA), Addis Ababa, Ethiopia. ECA/NRD/CART.9/ORG.3O Add.1, p. 8.

Cihlar J., 2000, Land cover mapping of large areas from satellites: status and research priorities. International Journal of Remote Sensing, 21(6&7): 1093–1114.

Di Gregorio A., 1991, Technical Report on Land Cover Mapping of Lebanon. FAO project document. Project NECP/LEB/001/SAU. Food and Agriculture Organisation of the United Nations (FAO), Rome, Italy.

Di Gregorio A., Jansen L.J.M., 1996, Africover Land Cover Classification. FAO Working Paper for the Dakar Meeting of Africover Working Group on Legend and Classification. Food and Agriculture Organisation of the United Nations (FAO). Rome, Italy, pp. 25+27 of appendices.

Di Gregorio A., Jansen L.J.M., 1996, The Africover Land Cover Classification Scheme: A Dichotomous, Modular-Hierarchical Approach. Working Paper with the Proposal for the International Working Group Meeting, 29–31 July, Dakar.

Ehlers M., 1995, The Promise of Remote Sensing for Land Cover Monitoring and Modeling, Joint European Conference and Exhibition on Geographical Information. The Hague, The Netherlands, 2, pp. 426–432.

European Environmental Agency, 2002, European Topic Center-Terrestrial Environment. CORINE Land Cover Update I&CLC2000 Project. Technical Guidelines.

European Environment Agency, 2006, The thematic accuracy of Corine Land Cover 2000 Assessment using LUCAS. EEA Technical Report No. 7/2006, ISSN 1725-2237.

European Environment Agency, 2007, Corine Land Cover 2000 (CLC2000) 100 m - version 9/2007.

EUROSTAT, 1999, Land Cover and Land Use Information Systems for European Union Policy Needs. Seminar Proceedings, Luxembourg. EUROSTAT, Studies and Research, Theme 5, p. 193.

FAO, 1994, Africover. Food and Agriculture Organisation of the United Nations (FAO). Project Document. FAO Remote Sensing Centre, Rome, Italy, p. 142.

FAO, 1997, Africover Land Cover Classification, Remote Sensing Centre Series, n.70, FAO, Rome.

FAO/UNEP/IAO/USAID, 2002, Proceedings of the FAO/UNEP Expert Consultation on Strategies for Global Land Cover. Mapping and Monitoring. Artimino, Florence, Italy, 6–8 May 2002.

Feranec J., Otahel J., Pravda J., 1995, Proposal for a Methodology and Nomenclature Scale 1:50,000 CORINE Land Cover Project, Final Report, Institute of Geography, Slovak Academy of Sciences, Bratislava.

Grossman D.H., Faber-Langendoen D., Weakley A.W., Anderson M., Bourgeron P., Crawford R., Goodin K., Landaal S., Metsler K., Patterson K.D., Pyne M., Reid M., Sneddon L., 1998, International Classification of Ecological Communities: Terrestrial Vegetation of the Unifed States, Vol. 1, The National Vegetation Classification System: Development, Status and Application. The Nature Conservancy, Arlington, Virginia, USA, p. 126.

Iliffe J., 1994, Africover - Geodesy and Map Projections. Unpublished FAO Background Paper for the Technical and Donor Consultations on the Africover Project, 4–11 July, Addis Ababa, Ethiopia.

Kalensky S.D., 1998, Africover land cover database and map of Africa. Canadian Journal of Remote Sensing, 24(3): 292–297.

Kalensky S.D., Cumani R., 2000, Use of Advanced Geo-Information Technologies in Land Use Planning and Management. Presented at the Geomatics Conference, Montreal, Canada, p. 12.

Kalensky S.D., Latham J.S., 1998, The Establishment of Environmental Information Systems (EIS) in developing countries. Geomatica, 52(4): 474–480.

Lambin E.F., Baulies X., Bockstael N., Fischer G., Krug T., Leemans R., Moran E.F., Latham J.S., He C., Alinovi L., Di Gregorio A., Kalensky S.D., 2002, FAO Methodologies for Land Cover Classification and Mapping. Chapter of the book: Remote Sensing and GIS Applications for Linking People, Place and Policy, S.J. Walsh and K.A. Crews-Meyer (Eds.) Kluwer Academic Publishers, Boston, USA.

Loveland T., Yasuoka Y., Burgan B., Chen J., Defries R., Lund H.G., Lynham T., Mayaux P., Gregoire J.-M., 1998, GOLD-3: Global Observations of Forest Cover: Coarse resolution Product Design Strategy workshop report, Sioux Falls, USA, July.

Rindfuss R.R., Sat Y., Skole D., Turner II B.L., Vogel C., 1999, Land Use Cover Change Implementation Strategy. IGBP Report N.48, IHDP Report: 10, Stockholm, International Geosphere-Biosphere Programme Secretariat.

Palko S., St-Laurent L., Huffman T., Unrau E., 1996, The Canada Vegetation and Land Cover: A Raster and Vector Data Set for GIS Applications – Uses in Agriculture. Published in GIS Applications in Natural Resources 2. M. Heit, H.D. Parker and A. Shortreid, Eds. GIS World Inc. Fort Collins, Colorado, USA, pp. 185–191.

Skole D.L., Salas W.A., Taylor V., 1998, GOLD-4: Global Observation of Forest Cover: Fine Resolution Data and Product Design Strategy Workshop Report, Paris, France, 23–25 September.

Thompson M., 1996, A standard land-cover classification scheme for remote sensing applications in South Africa. South African Journal of Science, 92(1): 34–42.

UNEP 1993, Vegetation Classification. Report of the UNEP-HEM/WCMC/GCTE Preparatory Meeting, Charlottesville, Virginia, USA. GEMS Report Series Publication, 19, p. 21.

Vogelmann J.E., Sohl T.L., Campbell P.E., Shaw D.M., 1998, Regional land cover characterisation using Landsat thematic mapper data and ancillary data sources. Environmental Monitoring and Assessment, 51: 415–428.

Vogelmann J. et al., 2001, Completion of the 1990s national land cover data set for the conterminous United States from Landsat thematic mapper data and ancillary data sources. Photogrammetric Engineering and Remote Sensing, 67: 650–652.

Young A., 1994, Towards International Classifi cation Systems for Land Use and Land Cover, UNEP/FAO Report of the Export Meeting on Harmonising Land Cover And Land Use Classifications, Geneva, 23–25 November, GEMS Report Series 25, UNEP, Nairobi, Kenya, p. 45.

Wyatt B.K. et al., 1997, Guidelines for Land Use and Land Cover Description and Classifi cation. Draft Final Report. ITE Huntingdon/ITC Enschede/WCMC Cambridge/UNDP Nairobi, p. 134.

Colour Plates

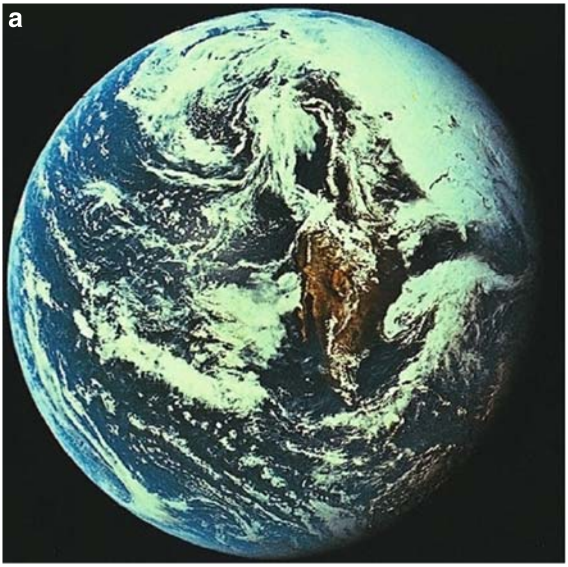
Plate 2.1 (**a**) Representation of the Earth as a cone projection with meridian and parallel, elaborated by Claudio Tolomeo (II century DC). (**b**) The Earth in a Tolomeo map, in *De geographia Latin Code*, XV century

M.A. Gomarasca, *Basics of Geomatics*, DOI 10.1007/978-1-4020-9014-1_BM2,

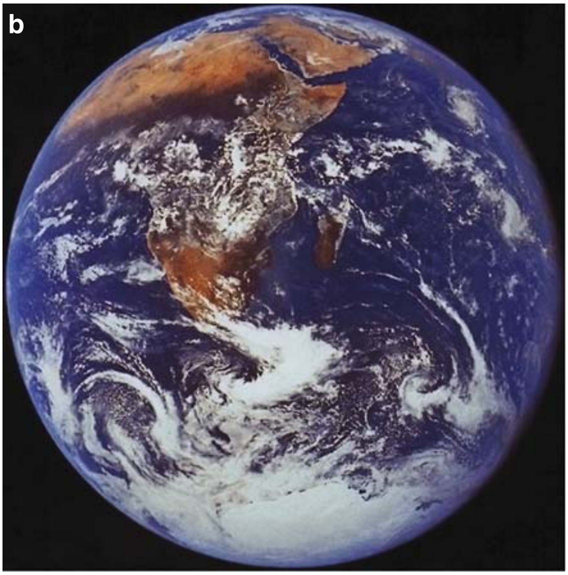
Plate 2.1 (continued)

Plate 2.2 The Piri Reis map edited in 1533 is considered probably the first and for sure the most precise document that represent the Americas in the XVI century

Plate 2.3 Earth from space by the Apollo 10 (**a**) and 17 (**b**) recorded in May 18, 1969 and December 7, 1972; the Apollo missions transported the man on the Moon July 20, 1969 with Apollo 11



Apollo 10 May 18, 1969



Apollo 17 December 7, 1972

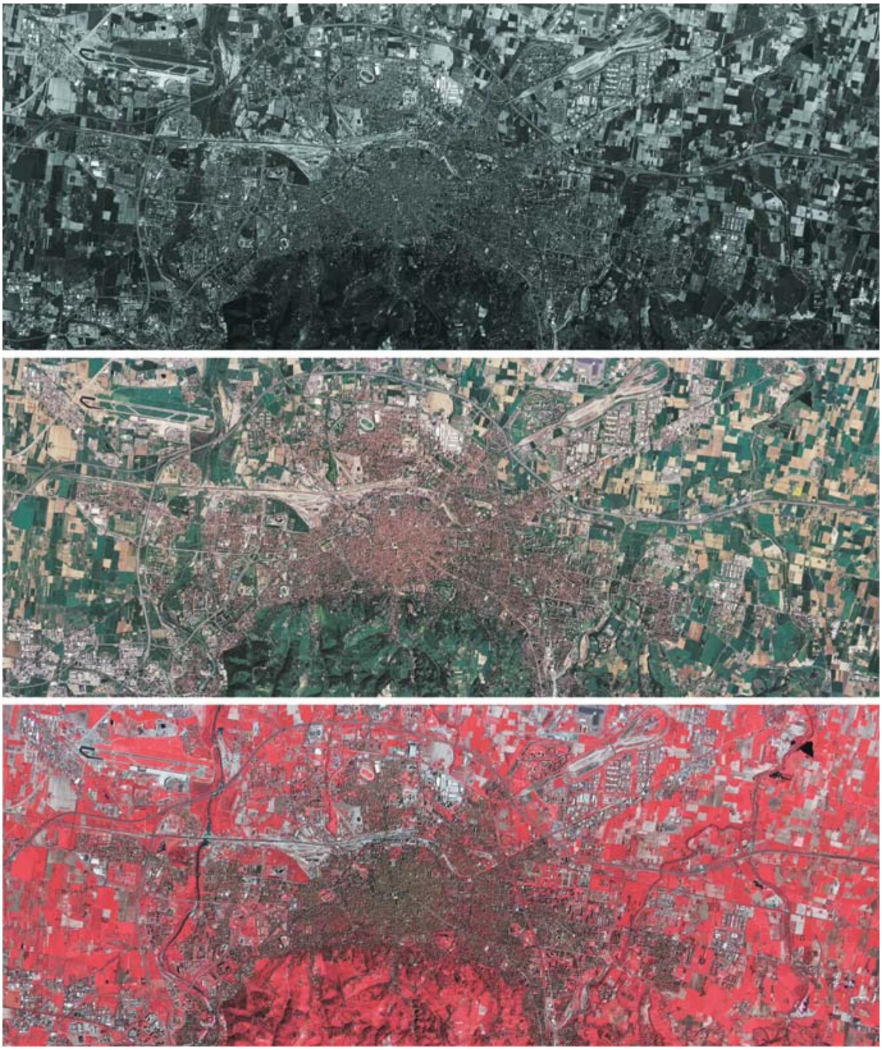


Plate 3.1 Aerophotogrammetric digital camera ADS40 with pushbroom linear sensor; (**a**) panchromatic, (**b**) true colours, (**c**) infrared false colours (© CGR, Parma). Relative altitude: 6.240 m. Flight data: 22nd of April 2004. Geometric resolution: 65 cm. Swath width: 7,8 km (12000 pixel × 65 cm). Swath length: 25 Km

a

b

Plate 3.2 Particular of and image recorded by the aerophotogrammetric digital camera ADS40. (**a**) colour, (**b**) infrared false colour (© CGR, Parma)

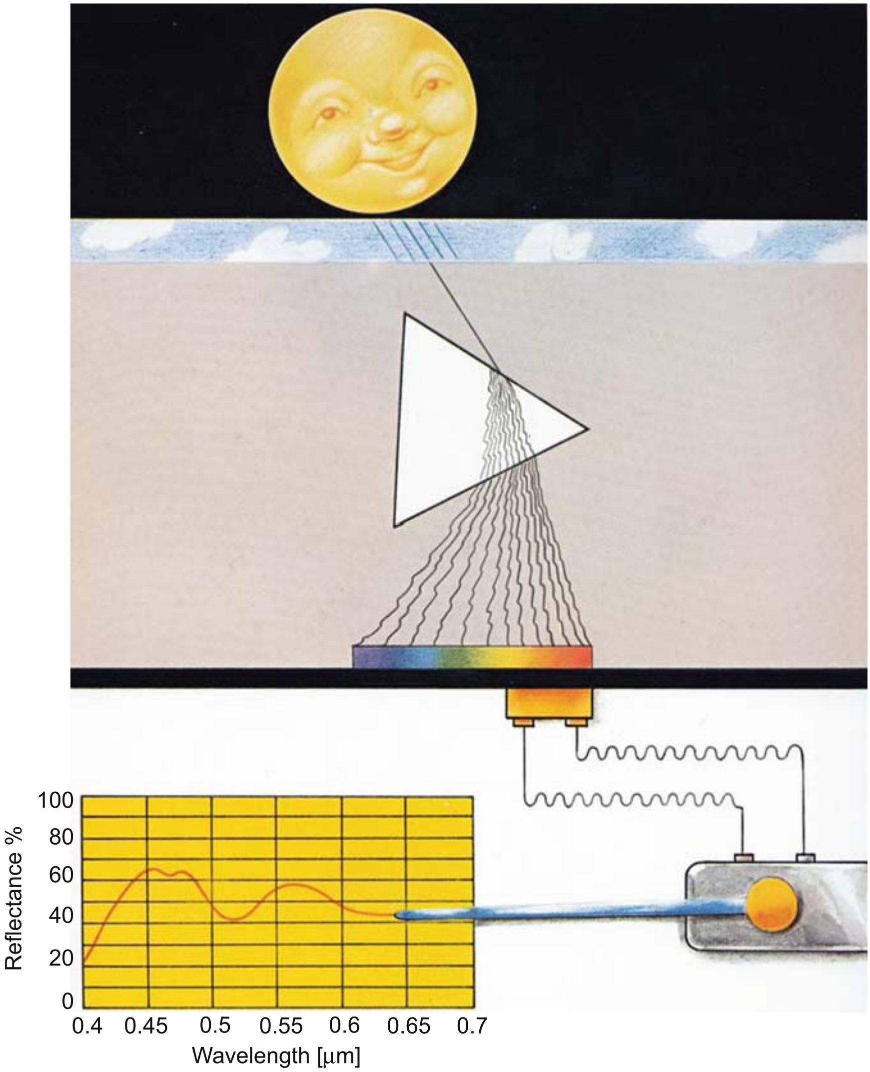


Plate 4.1 The visible interval (0.38 – 0.75 μm) of the electromagnetic spectrum passing through a prism is split in the rainbow colours from the violet (0.40 – 0.41 μm) to the red (0.65 – 0.68 μm), as experimented by Newton in 1666. Max Plank in 1900 has drawn the bases to measure the intensity of each colour of the visible light

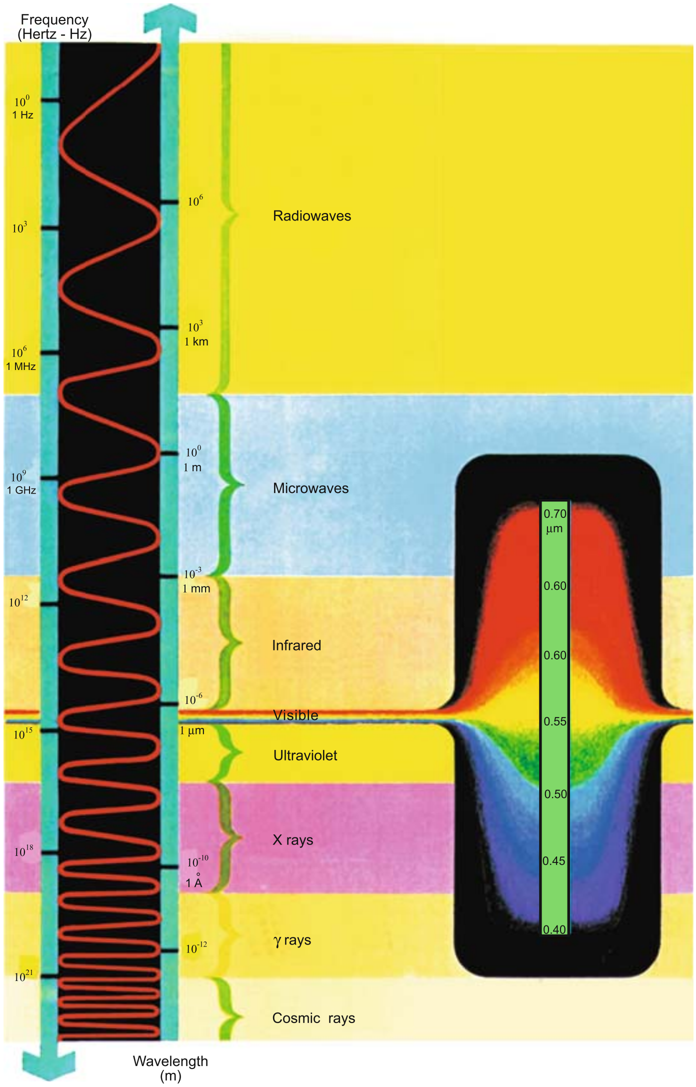


Plate 4.2 The electromagnetic spectrum subdivided in its characteristics regions, expressed by frequency (Φ:Hz) and wavelength (λ: µm). The wavelength is the inverse of the frequency

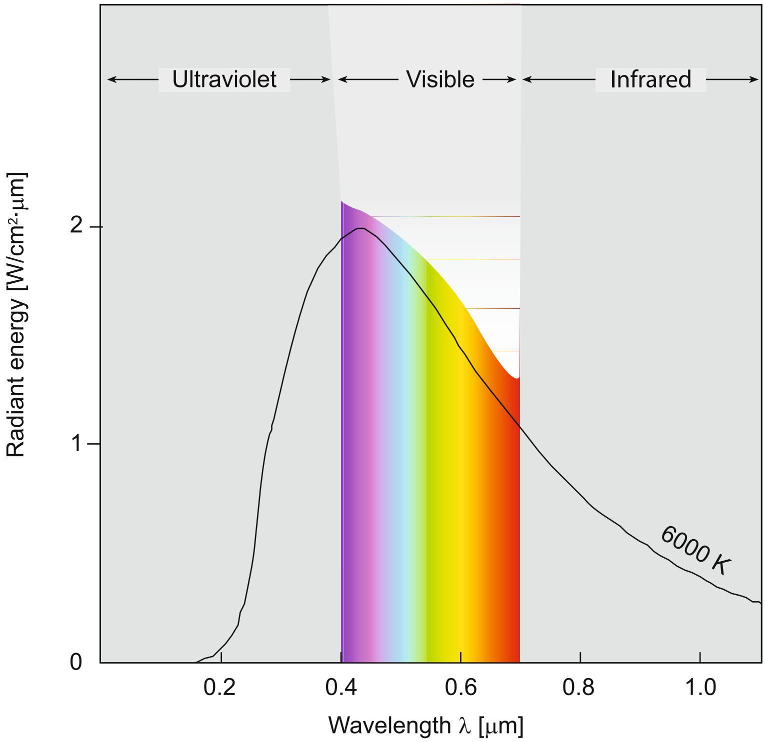


Plate 4.3 The region of the visible from 0.4 to 0.7 μm of wavelength is subdivided in the seven fundamental colours; starting from the shorter λ are the following: *violet*, *blue*, *cyan*, *green*, *yellow*, *orange*, *red*

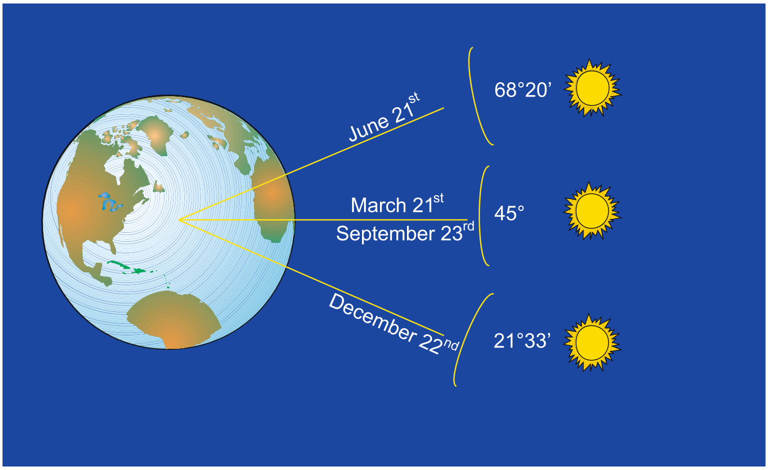


Plate 4.4 Sun angle at noon in different seasons. At latitude of 40° – 45° the highest solar energy occurs the 21st of June, while 22nd of December there is the minimum availability of energy

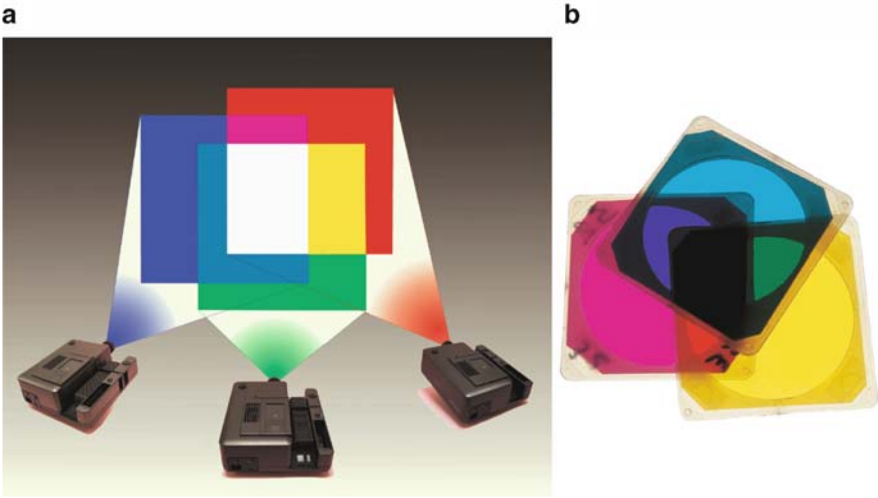


Plate 4.5 (**a**) additive synthesis of primary colours *Blue*, *Green* and *Red* projected on a white screen: their synthesis produce the *white* colour. The overlap of two primary colours generates the complementary colours *Yellow* (Y), *Magenta* (M) and *Cyan* (C); (**b**) subtractive synthesis, obtained by transparency starting form the white light; combining two of the three filters (M+C, Y+C, Y+M) the primary colour *Red*, *Green* and *Blue* are respectively transmitted. The overlap of the three filters Y M C determines the absorption of the three colours of the white light, resulting in the black

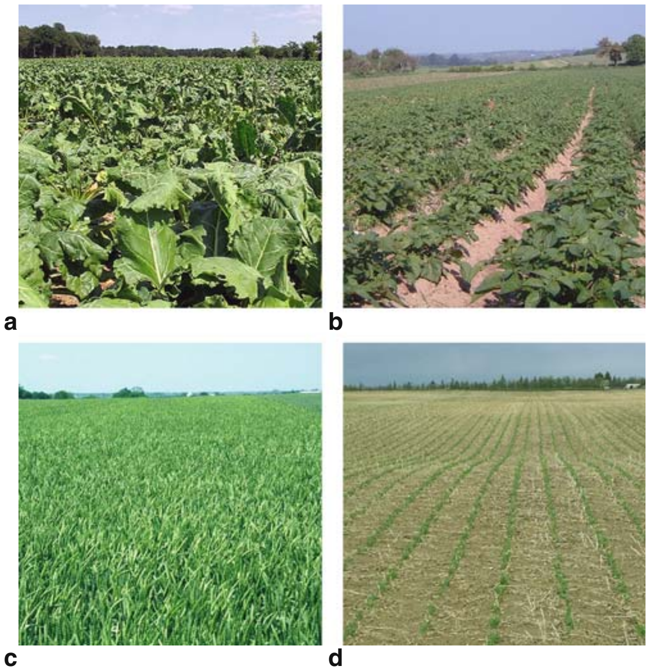


Plate 4.6 Agriculture texture of some crops generating different radar backscattering signals; (**a**) sugar beet, (**b**) potato, (**c**) wheat; (**d**) beans

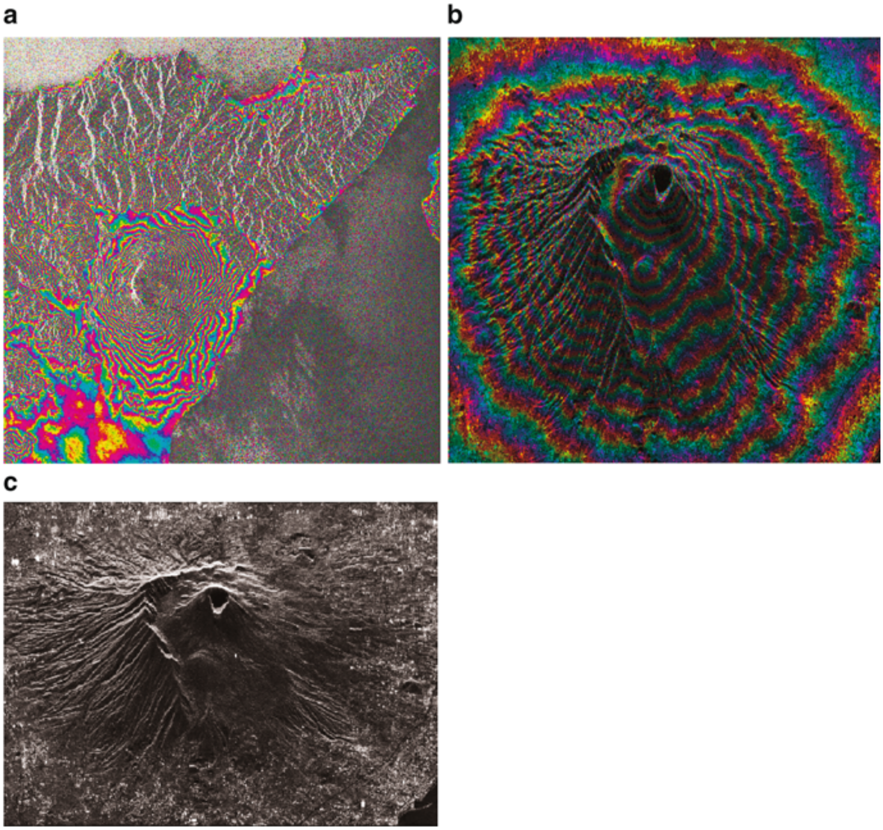


Plate 4.7 Interferometric phase in colours and module represented with the intensity of SAR ERS-1 images. The interferometric phase has been obtained as difference of two ERS-1 and ERS-2 images respectively on September 5 and 6, 1995. The Interferometric franges reproduce the contour lines; (**a**) the Etna volcano in Sicily with baseline ∼110 m; (**b**) Vesuvio volcano with baseline ∼135 m; (**c**) reference Landsat image (© DEI, Politecnico di Milano)

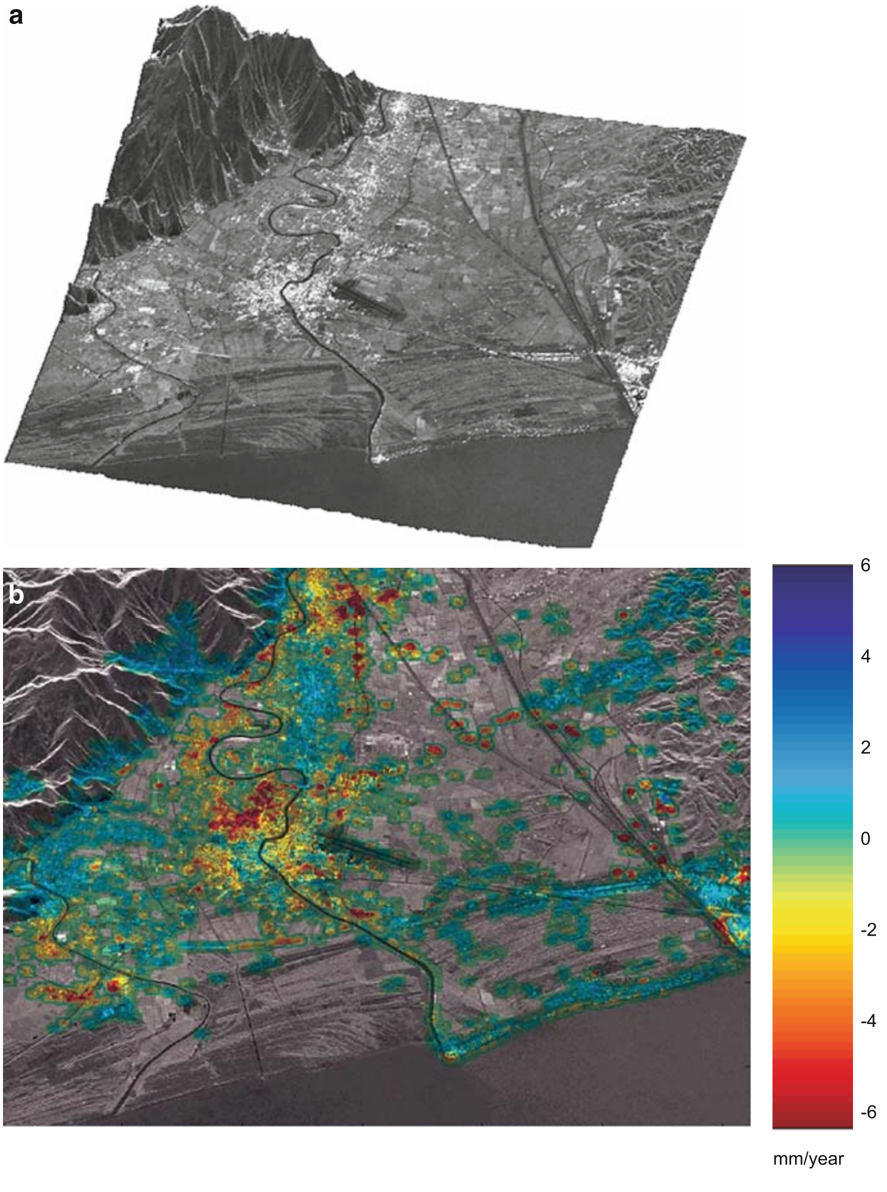


Plate 4.8 Mean velocity of land displacement along the view line in mm/year calculated with the permanent scatterers technique (PS); (**a**) 3D representation of the study area; (**b**) displacement in the time based on the reference scale (© DEI, Politecnico di Milano)

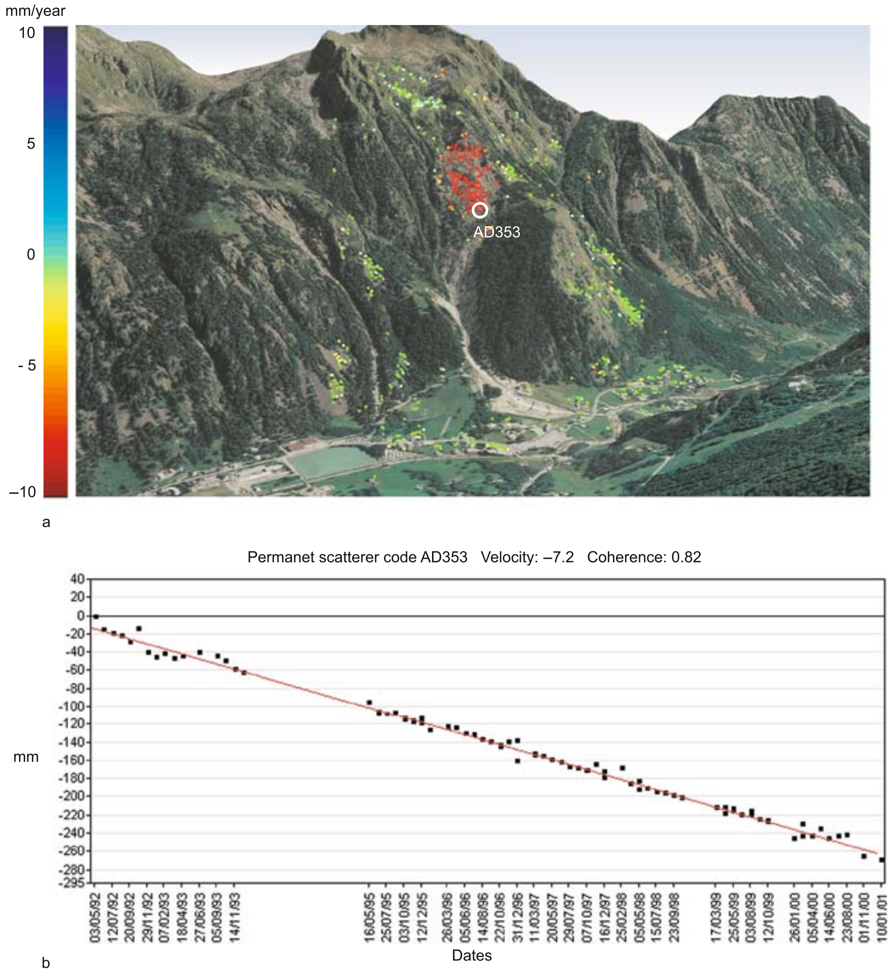
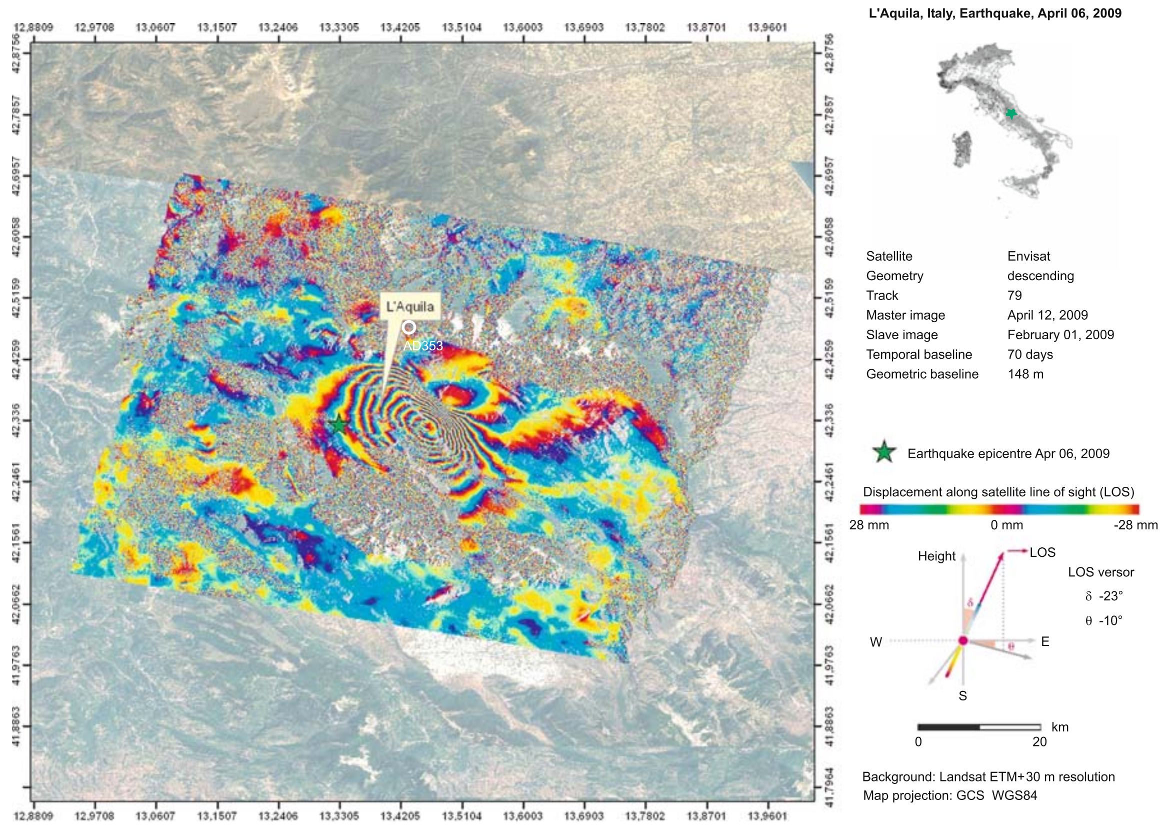


Plate 4.9 (**a**) 3D view of Bosmatto landslide, northern Italy: elaboration with the permanent scatterers technique (PS) of ERS data in the period 1992–2000; in the three-dimensional image is reported the mean velocity of deformation of the radar bench mark (PS) present in the study area. The deformation velocity of the PS are saturated in the range –10 (red) +10 (blue) mm/year. On the landslide slope are discernable several PS. Background image: Orthophoto + DEM 10 m (Tele-Rilevamento Europa, Milan). (**b**) Historic series of deformation of the permanent scatterer PS: AD353 (see Plate 4.9a)



b

Dates

Plate 4.10 L'Aquila, Italy Earthquake April 6, 2009. Co-seismic interferogram from ENVISAT data. The two images were acquired on 01/02/2009 and 12/04/2009. A co-seismic interferogram is a comparison of two radar images: one taken before the event and one after. The resulting interferogram shows where surface deformation caused by the earthquake was most significant, as marked by the fringes of colour. The accurate assessment of fault displacement and orientation (seismograms) can provide valuable input to seismologists for modeling the earthquake mechanism. (Courtesy: Tele-Rilevamento Europa, Milan)

Plate 6.1 The principle of acquisition of the remote sensing imagery

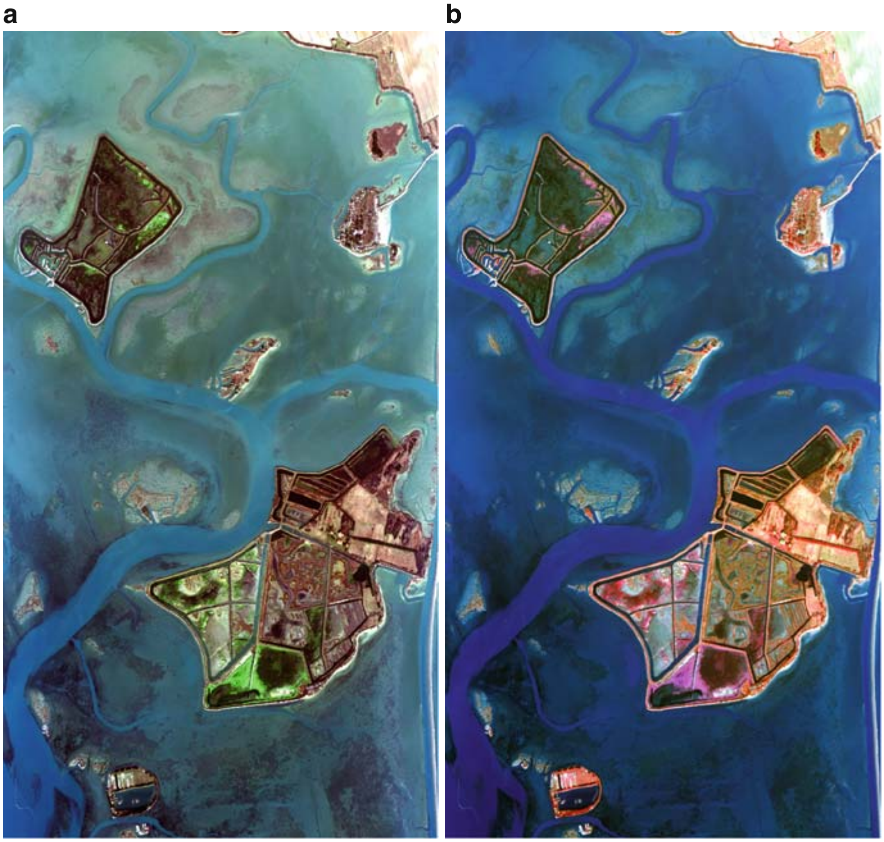
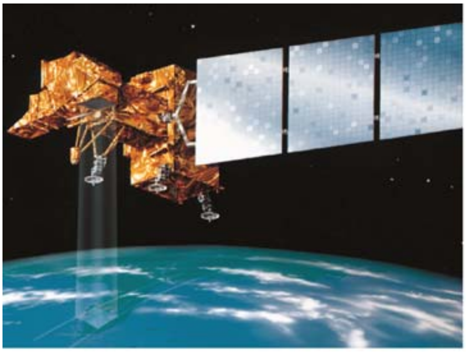
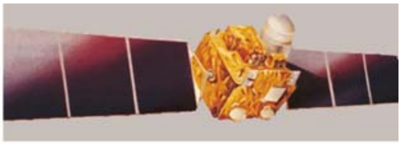


Plate 6.2 Two images of the same area in true colour, **a**, and false colour, **b**. In **b** the vegetation is represented by hue of *red* and *magenta*. The water, with very low reflection in the *red* and *infrared*, has *blue-black* colour. Grado Lagoon, Italy





IRS-1C

Landsat

ERS-1

Envisat

QuickBird

Plate 6.3 Payloads of some satellites orbiting the Earth: Landsat, IRS-1C, ERS-1 twin of ERS-2, Envisat and QuickBird

Plate 6.4 The configuration of the International Space Station (ISS)

Plate 6.5 The line day/night of the sunset in Europe and Atlantic Africa recorded from the Space Shuttle

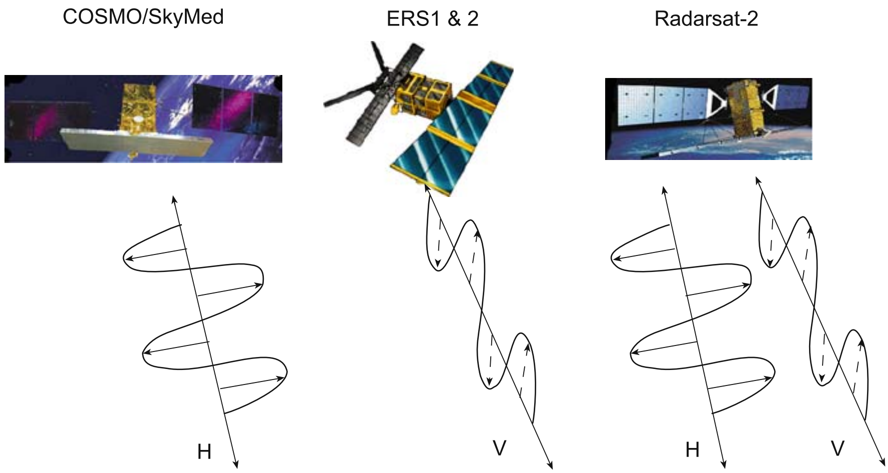


Plate 6.6 Types of polarization of the radar signal of some active satellite systems

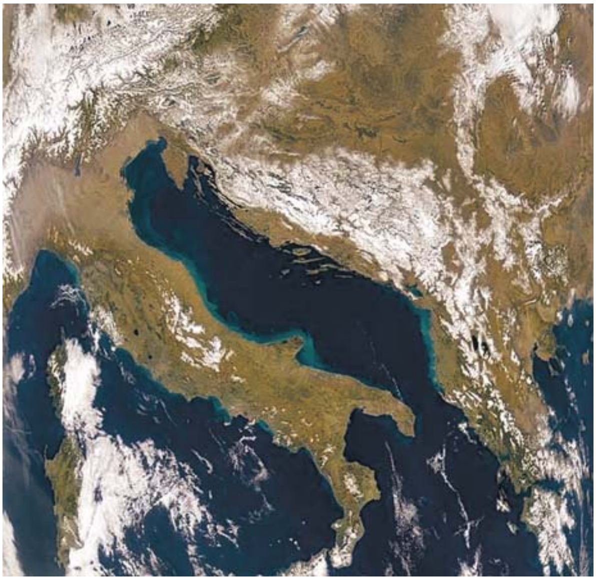


Plate 6.7 The sensor SeaWiFS on board the satellite OrbView-2; image of February 27, 1999 covering Italy and the Balkan area

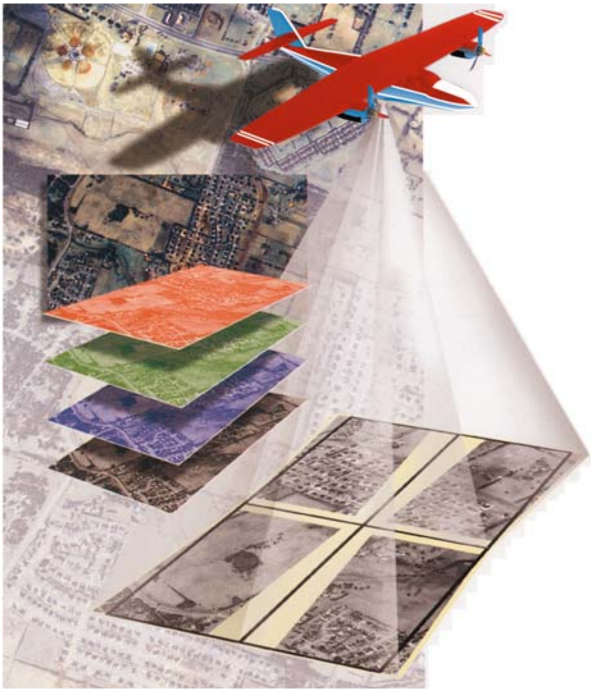
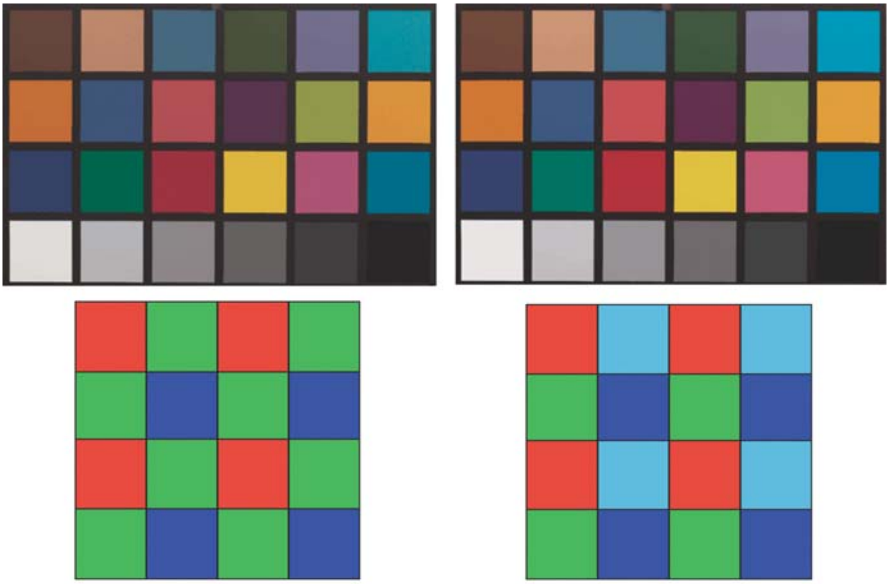


Plate 6.8 The digital camera DMC has different sensors: 3 central elements for the acquisition of panchromatic and/or colour imagery, 4 lateral elements for the multispectral acquisition (*blue*, *green*, *red*, near *infrared*). The simultaneous acquisition of 4 sub-scenes requires a mosaic reconstruction of the entire scene



a b

Plate 6.9 Two reference colour tables obtained by correction with two types of filter; (**a**) traditional RGB (*Red*, *Green*, and *Blue*); (**b**) four colour filter RGBE (*Red*, *Green*, *Blue*, and *Emerald*) that enlarges the available palette of colours improving the response and the chromatic variability

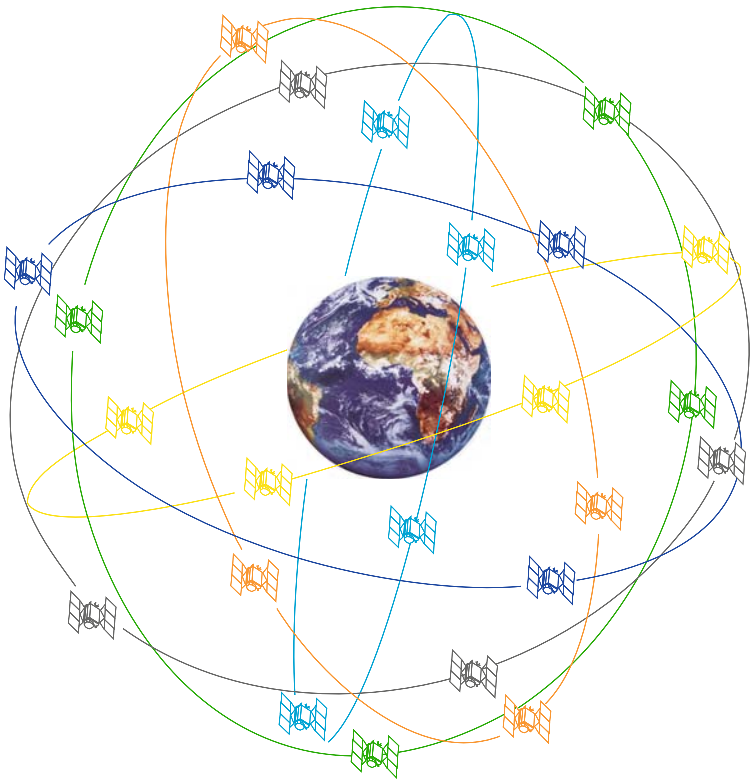


Plate 7.1 Constellation of the 24 satellites NAVSTAR distributed on 6 orbital planes. The orbits are quasi-polar at 20,183 km of altitude

Plate 7.2 European network EUREF of the permanent stations GPS

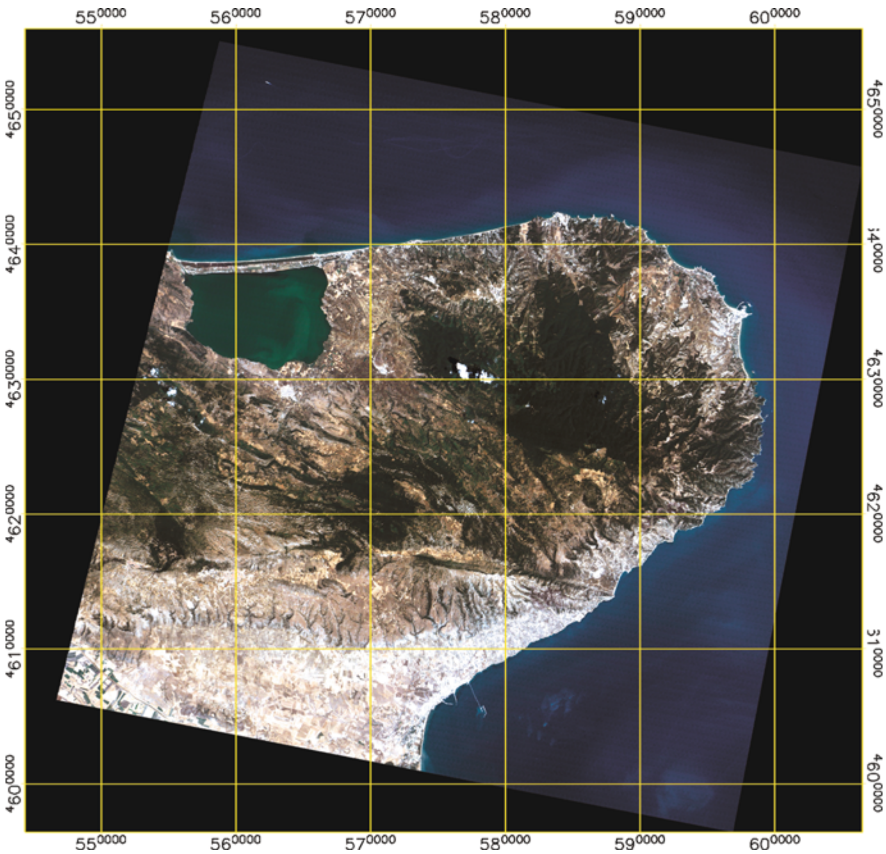


Plate 8.1 Landsat 5 TM image of the Gargano, Adriatic Sea, Italy, with kilometric grid of 10 km, UTM 32 N WGS84

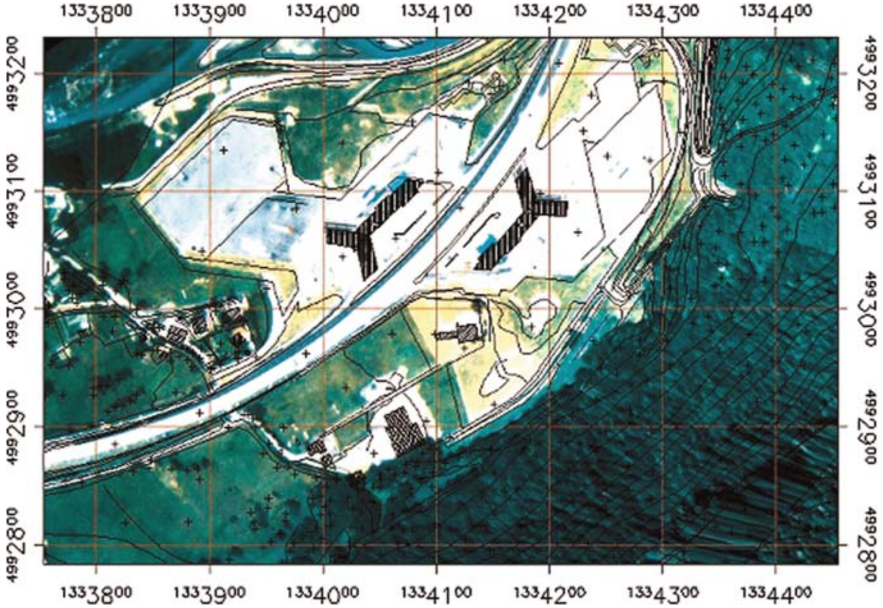


Plate 8.2 Overlap of a digital map and a projected image. The Root Mean Squared Error (RMSE) produced in the process of ortho-correction can be estimated observing an element both on the map (*black lines*) and on the digital image

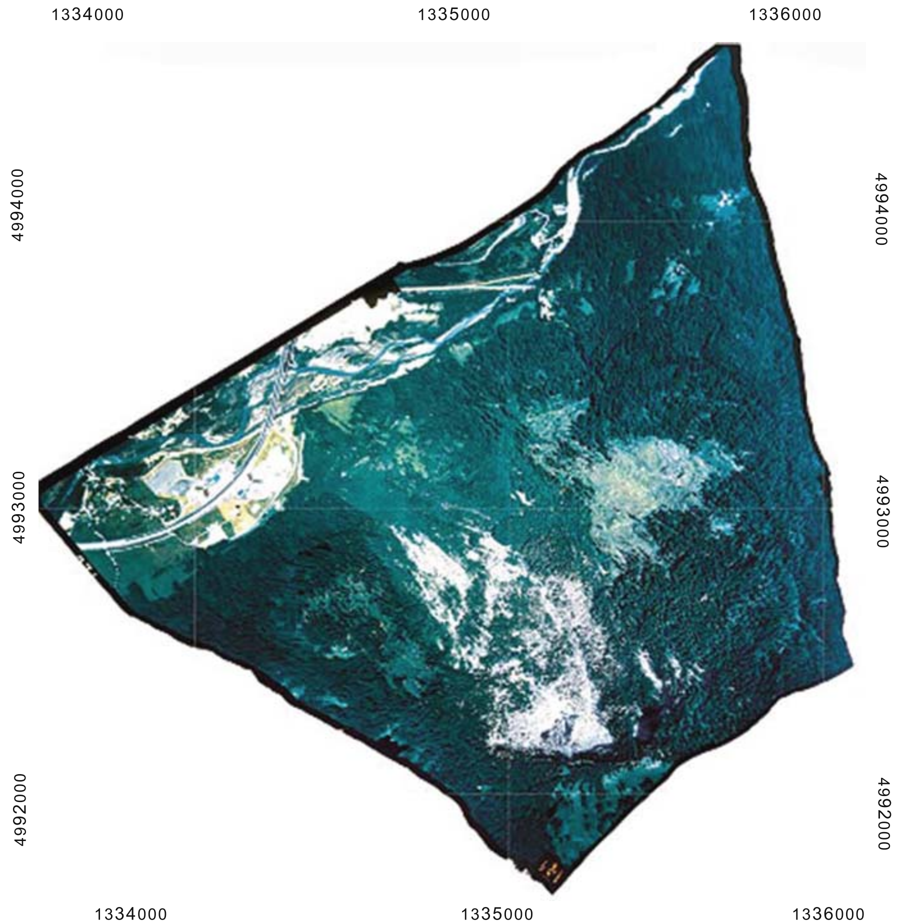


Plate 8.3 Ortho-projected photogram with original resolution of 1 m; the position of each element is defined both by image coordinates (path and row) and geographic coordinates (east, north)

Plate 8.4 Particular of the Plate 8.3

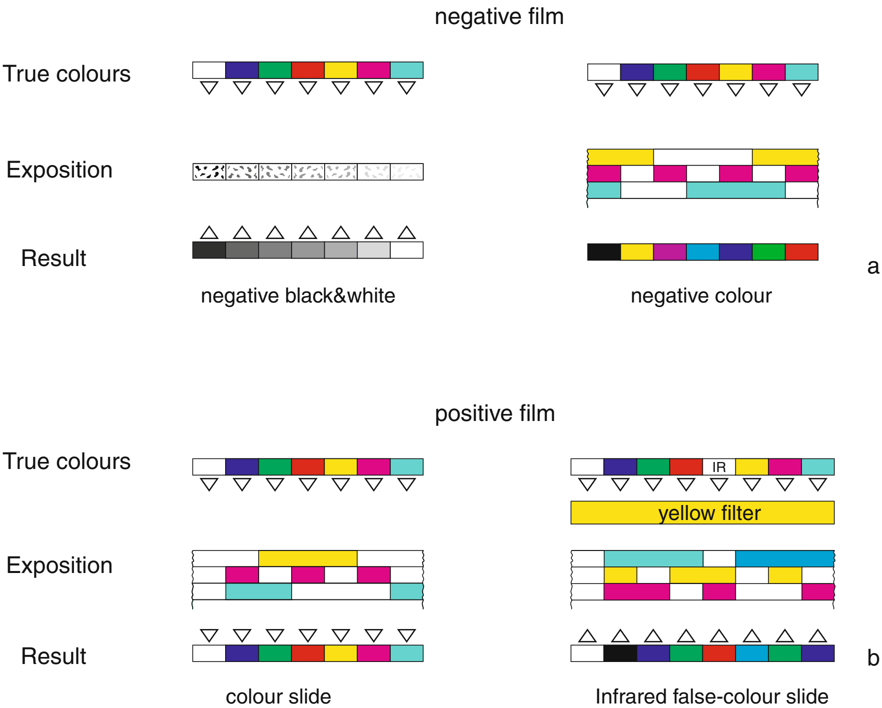


Plate 8.5 Structure and photographic process for the development of negative *black/white* and colour films (**a**) and slides (**b**). IRFC: InfraRed False Colour

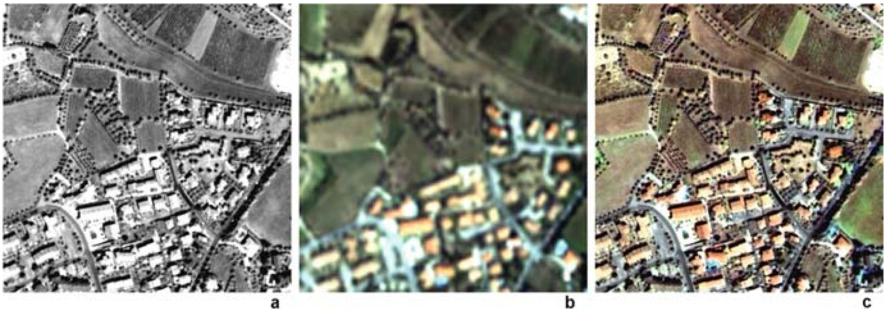


Plate 8.6 The fusion of two images by means of the pan-sharpening technique; a panchromatic image (better geometric resolution) is combined with the 3 multispectral bands (better spectral resolution) of the same scene obtaining a new synthetic image with enhanced geometric and spectral resolutions

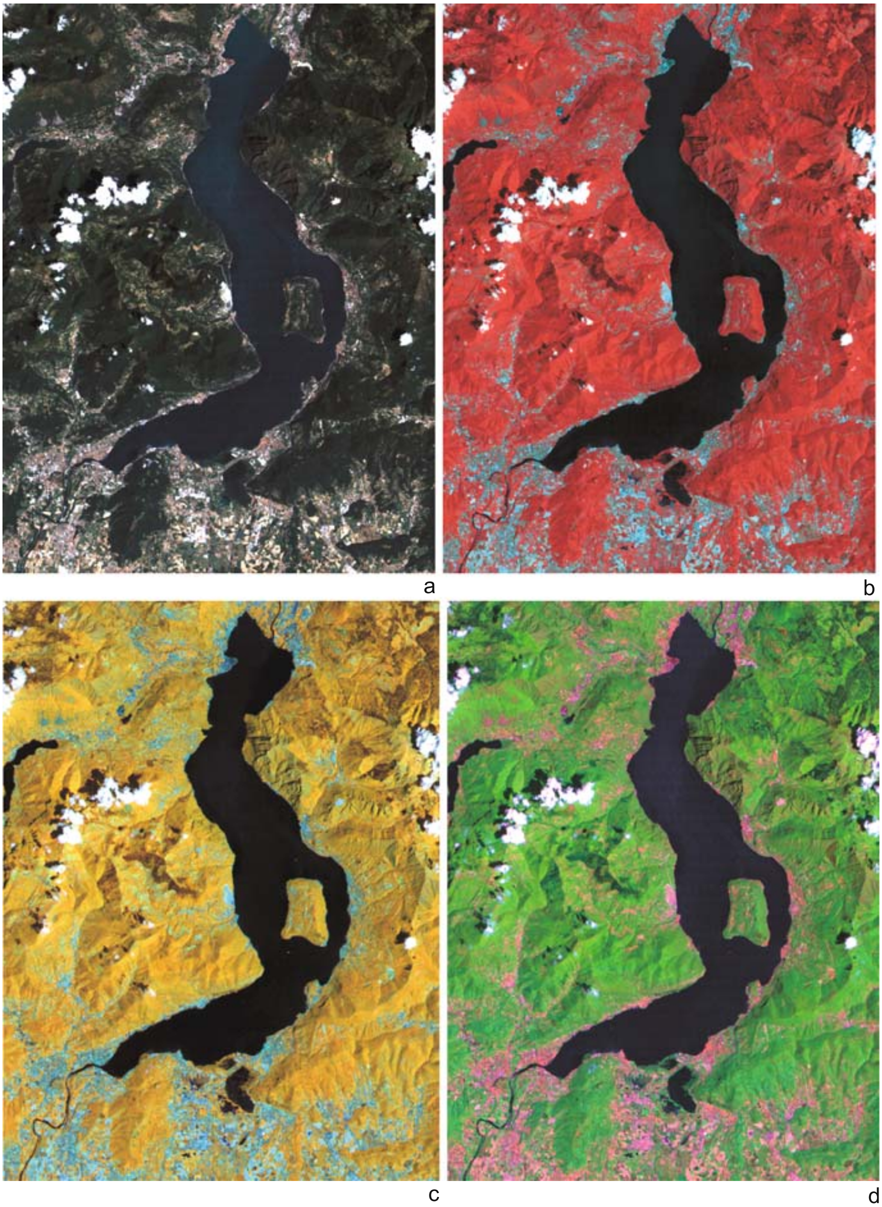


Plate 8.7 Landsat ETM + multispectral colour composites; (**a**) RGB: 321, (**b**) RGB: 432, (**c**) RGB: 453, (**d**) RGB: 741 (Iseo Lake, northern Italy)

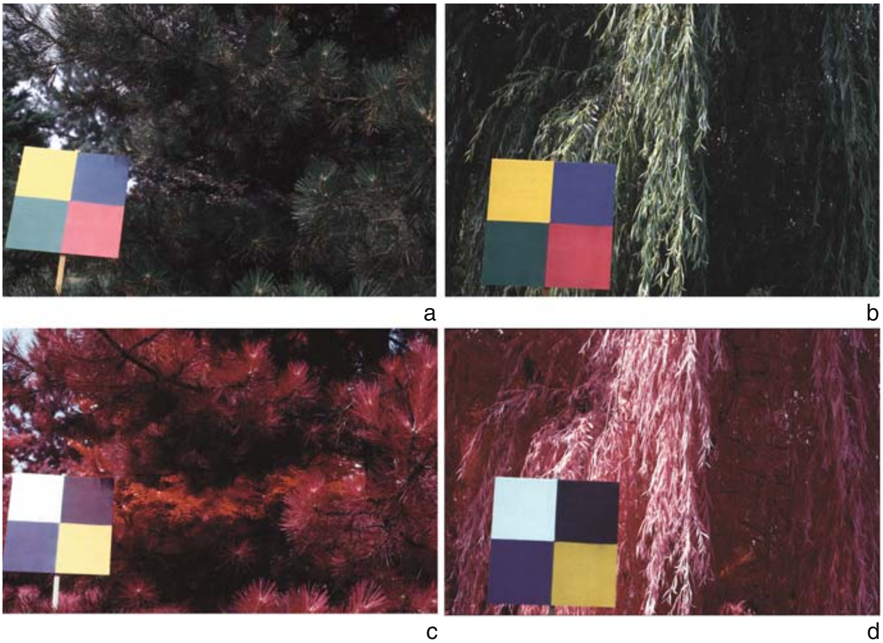


Plate 8.8 Colour (**a, b**) and infrared false colour (**c, d**) photographic films of coniferous (a) and broadleaf (b). In the reference panel the *green* colour in the colour film shot changes in the *blue* in the infrared false colour film (*yellow* filter on the lens tube) and the vegetation changes from *green* in *magenta* (due to the high reflectance of the vegetation in the NearIR)

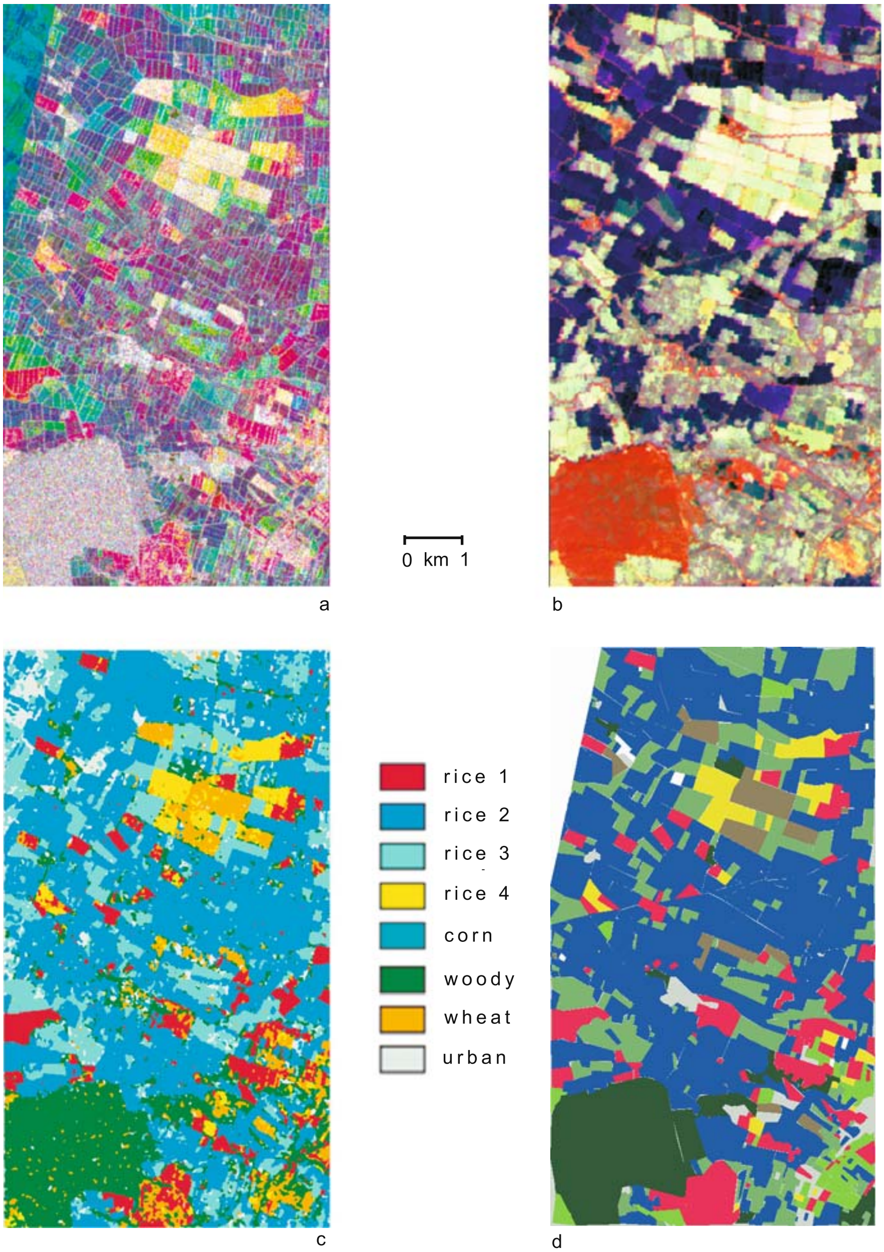


Plate 8.9 (**a**) multispectral colour composites of ERS-1 SAR, RGB: April 18, April 25, July 1, 1994; images compared with (**b**) Landsat RGB: 453 multispectral colour composites April 7, 1994 in a rice cultivated area. Pixel based (**c**) and field based (**d**) classifications of radar images

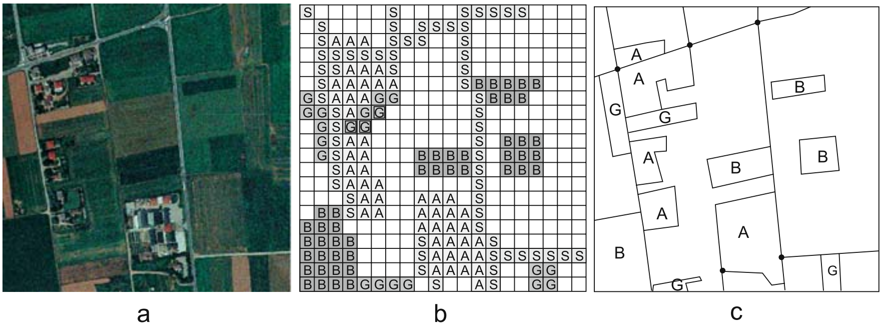


Plate 9.1 Vector and raster models: (**a**) natural colour from aero-photogram or digital multispectral sensor; (**b**) raster or grid model; (**c**) vector model

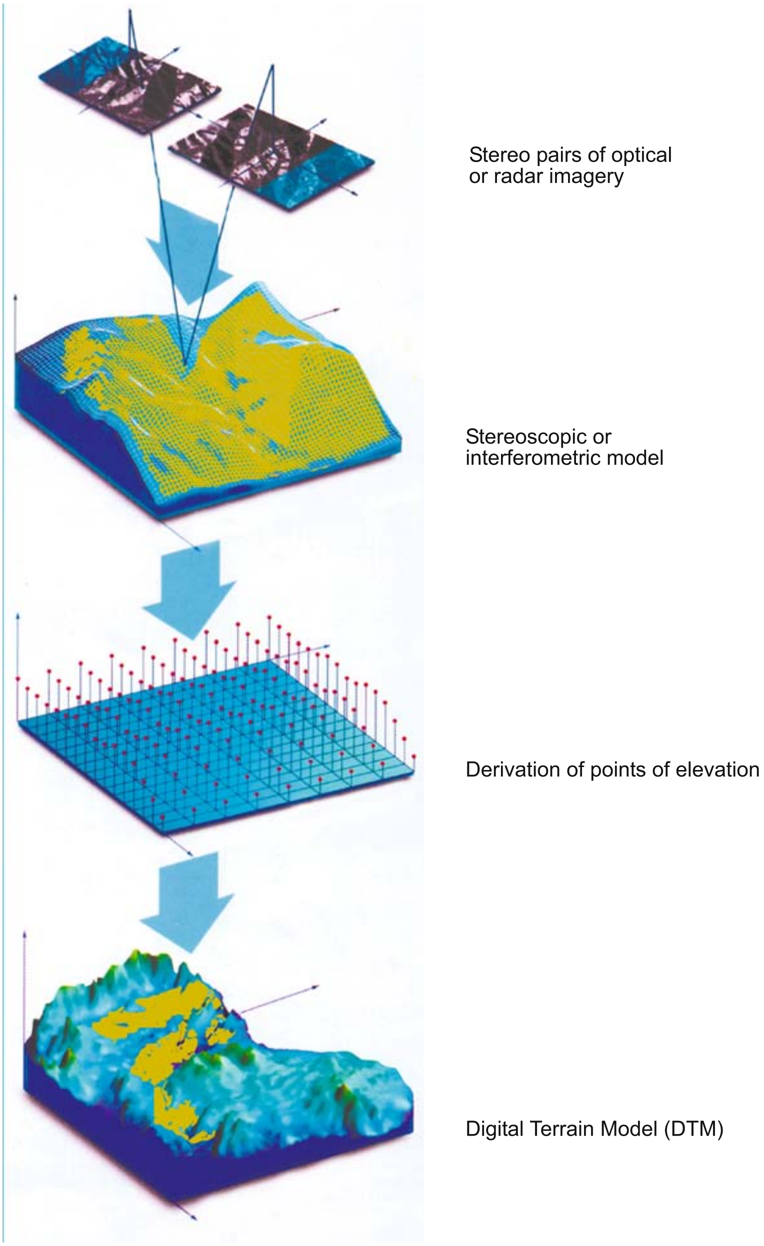


Plate 9.2 The process to obtain a Digital Terrain Model (DTM) from aero-photogrammetry or remote sensing data; from the stereoscopic model (derived from stereo-pairs or stereoscopic optical images) or the interferometric model (derived from radar images) are measured the contour lines or the interferometric fringes and the relative altitude of several points in the image are derived. By interpolation of the points a Digital Terrain Model (DTM) is represented

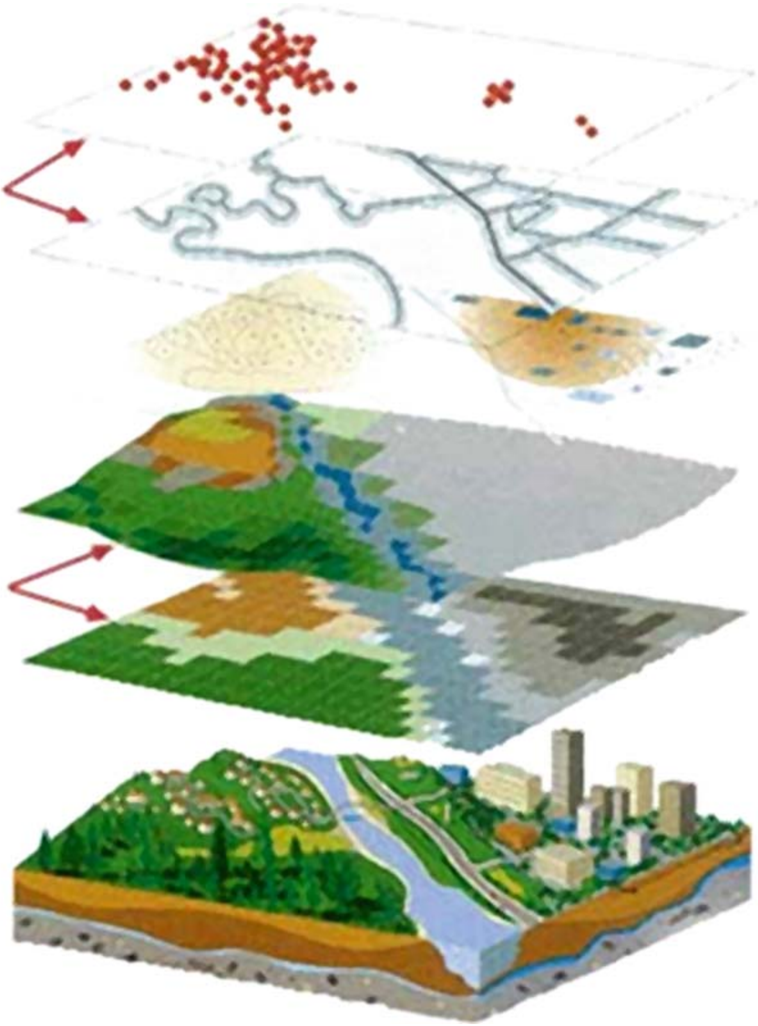


Plate 9.3 In the Geographical Information System the real world is stratified in several geocoded raster or vector information layers

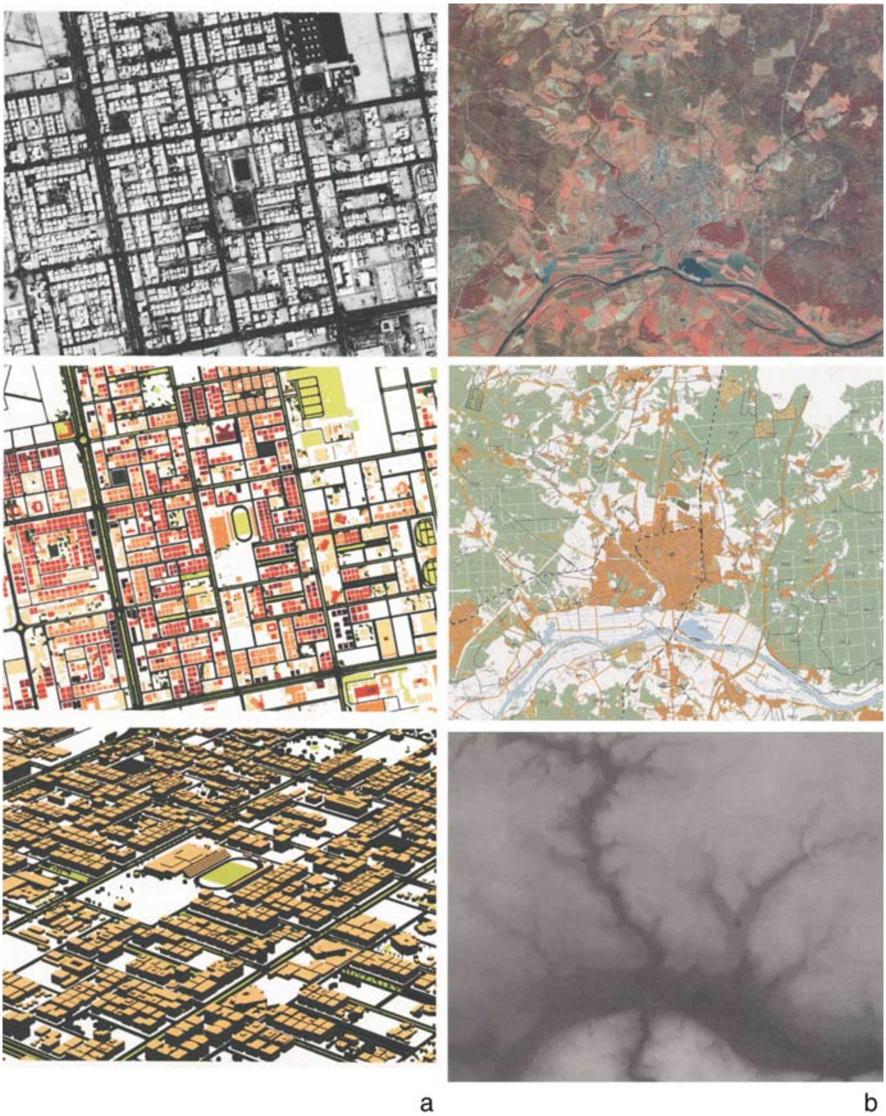


Plate 9.4 The 3D cartography is produced merging the information of the 2D cartography with the heights derived from aero-photogrammetric and/or satellite (optical and radar) acquisitions, and laser scanning systems

a

b

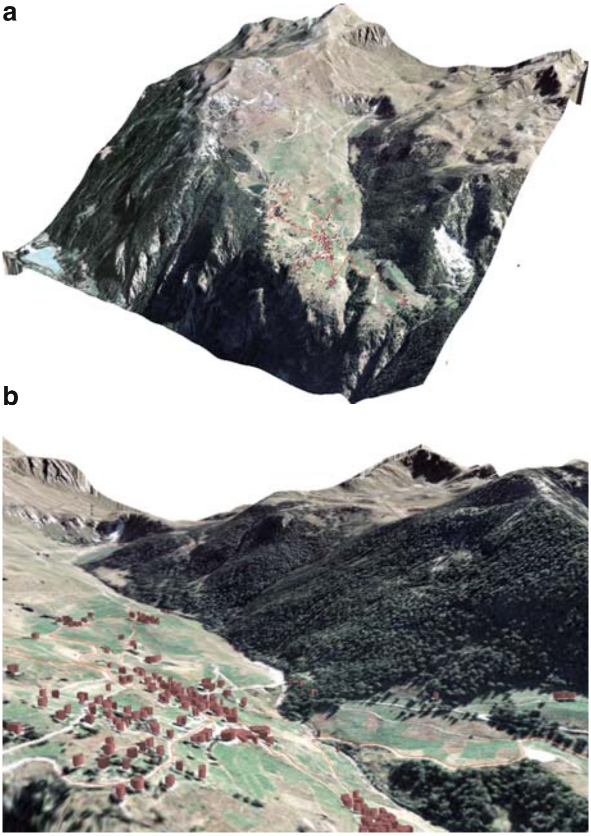


Plate 9.5 Model of a mountain landscape obtained combining a DEM, a digital orthophoto and an urban/building information layer; (**a**) full scene, (**b**) detail of the scene in (**a**)

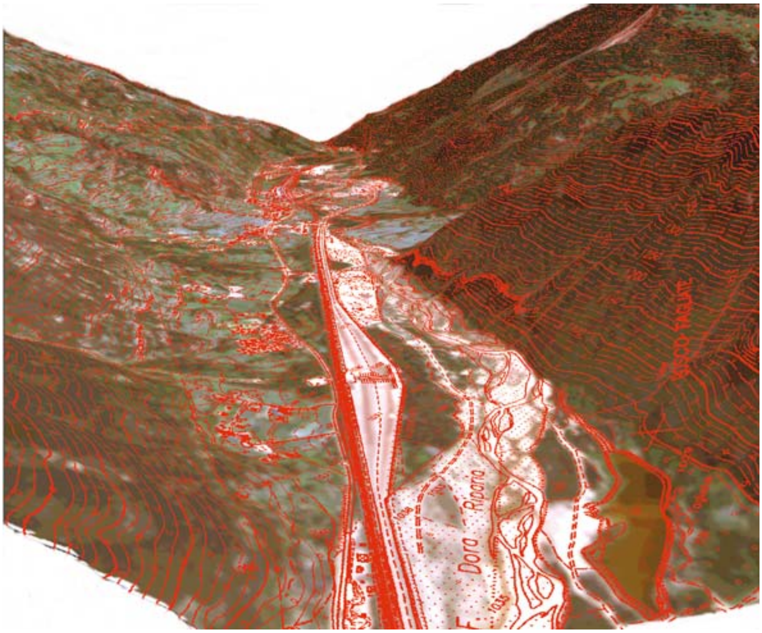


Plate 9.6 Hyperspectral MIVIS acquisition in true colour of a mountain region and the correspondent map, nominally in scale 1:10,000, combined with the DEM, step 50 m, of the area

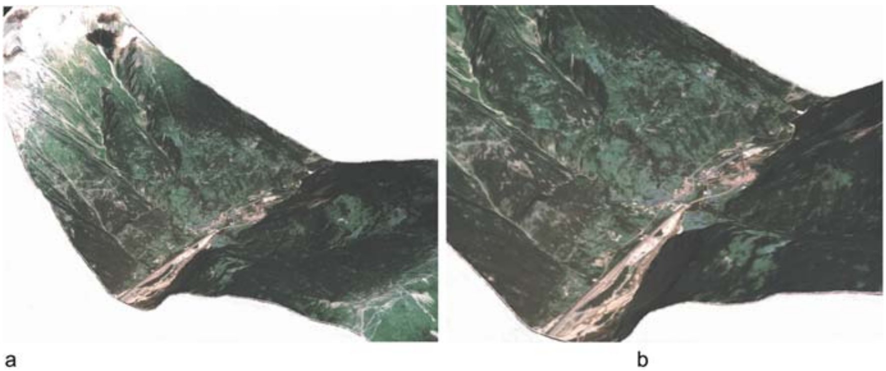


Plate 9.7 Hyperspectral MIVIS acquisition in true colour of a mountain region; (**a**) 45° E view of the scene in Plate 9.6, (**b**) zoom of (**a**)

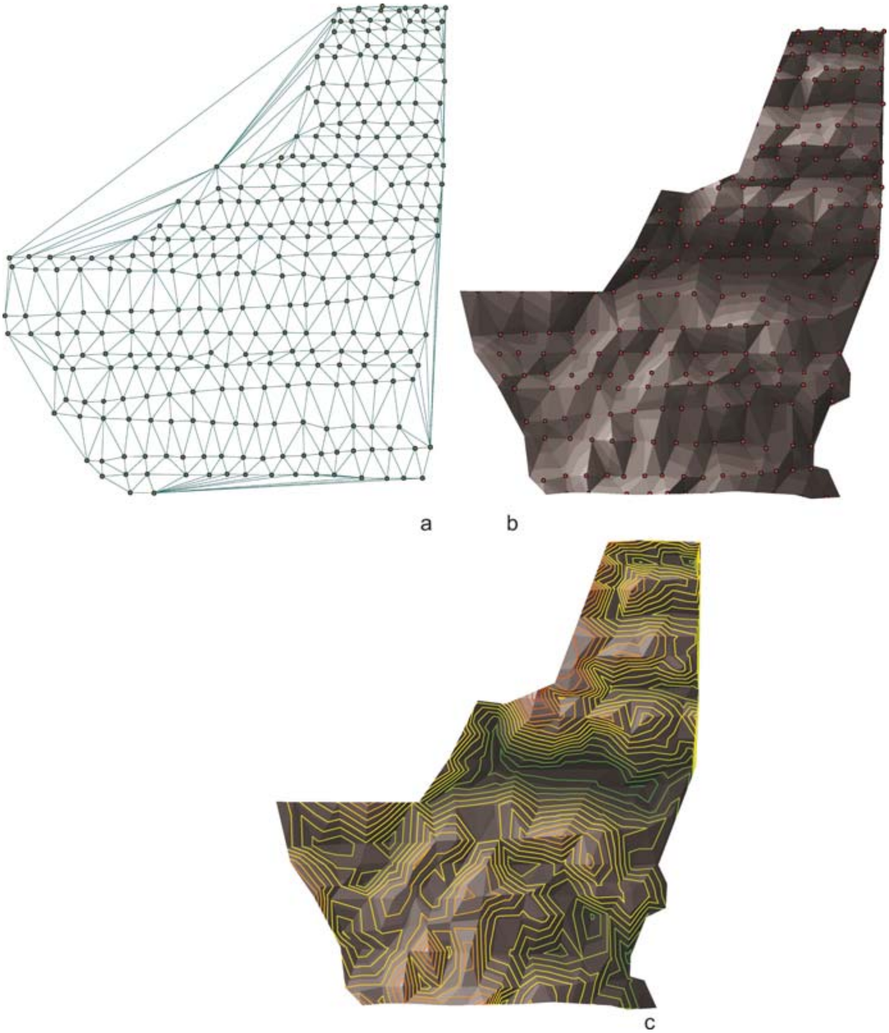


Plate 9.8 (**a**) TIN (Triangulated Irregular Network) derived from scattered points on two-dimensional plane based on Delaunay's triangulation. If the points have altitude information (z coordinates), generated TIN can be used for perspective viewing, (**b**) TIN with original scattered points overlap, (**c**) contour lines overlapping the TIN of generation. This data structure allows data to be displayed as three-dimensional surface, or to be used for terrain analysis including contouring and visibility mapping

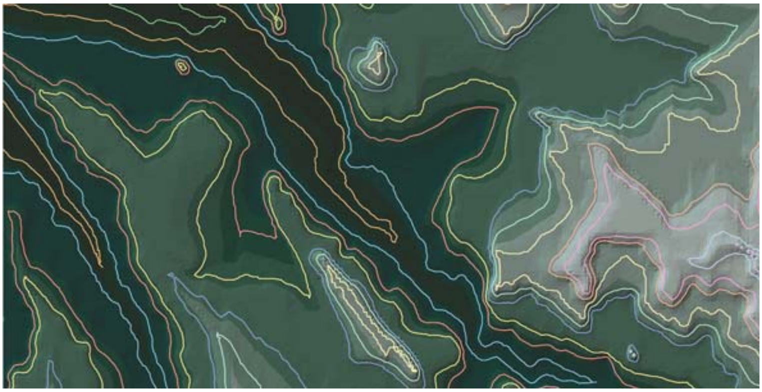


Plate 9.9 Contour lines overlapping the TIN of generation (see Fig. 9.29)

Plate 10.1 Example of ancillary data used for the realization of the CORINE program; (**a**) topographic map 1:25,000, (**b**) Land use/land cover map, (**c**) aerial photogram 1:50,000, (**d**): satellite image colour composite with overlay of the land use map

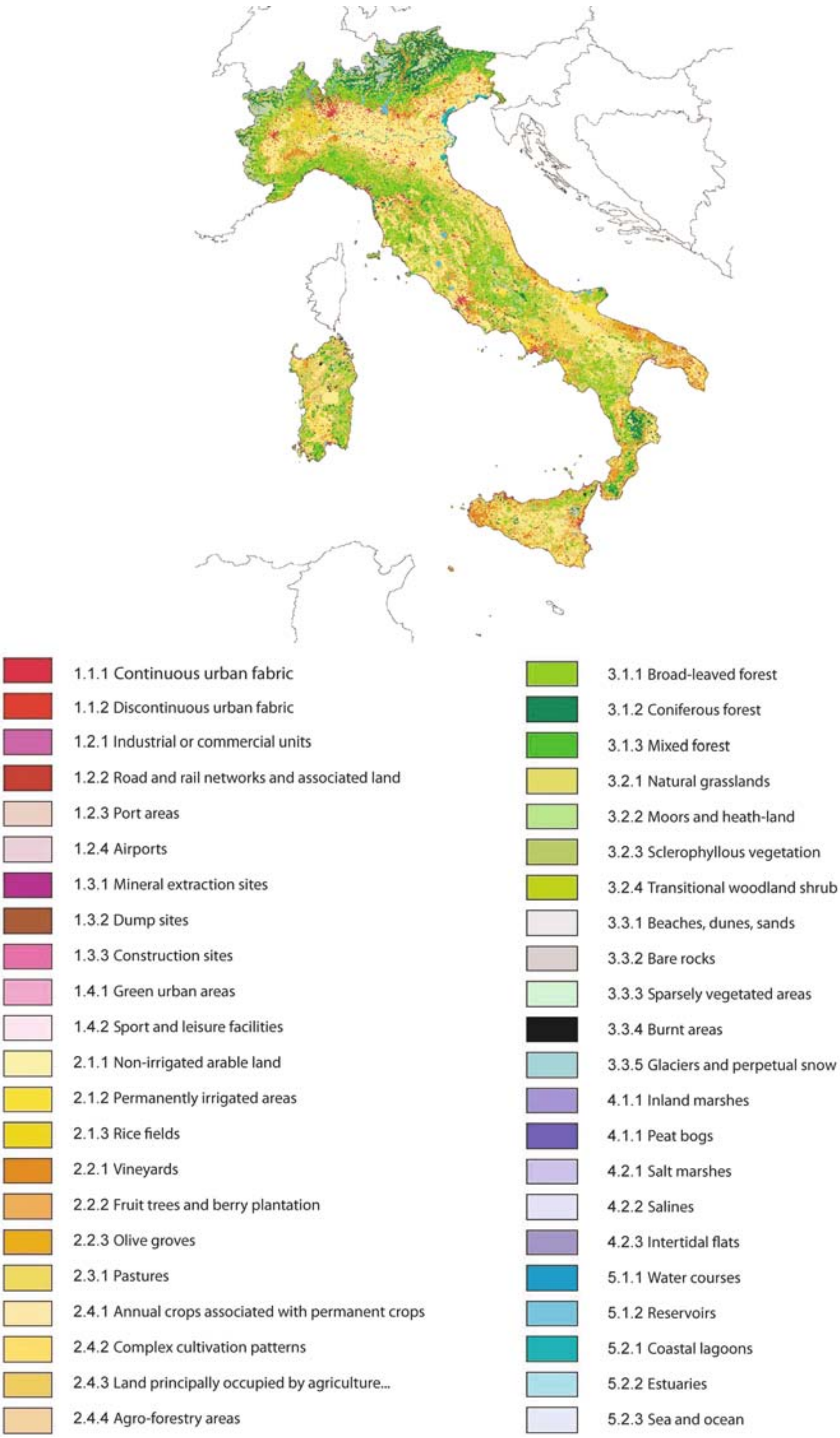


Plate 10.2 (**a**) Map of the CORINE-Land Cover classification of Italy; (**b**) Nomenclature in 44 classes of the CORINE-Land cover classification system

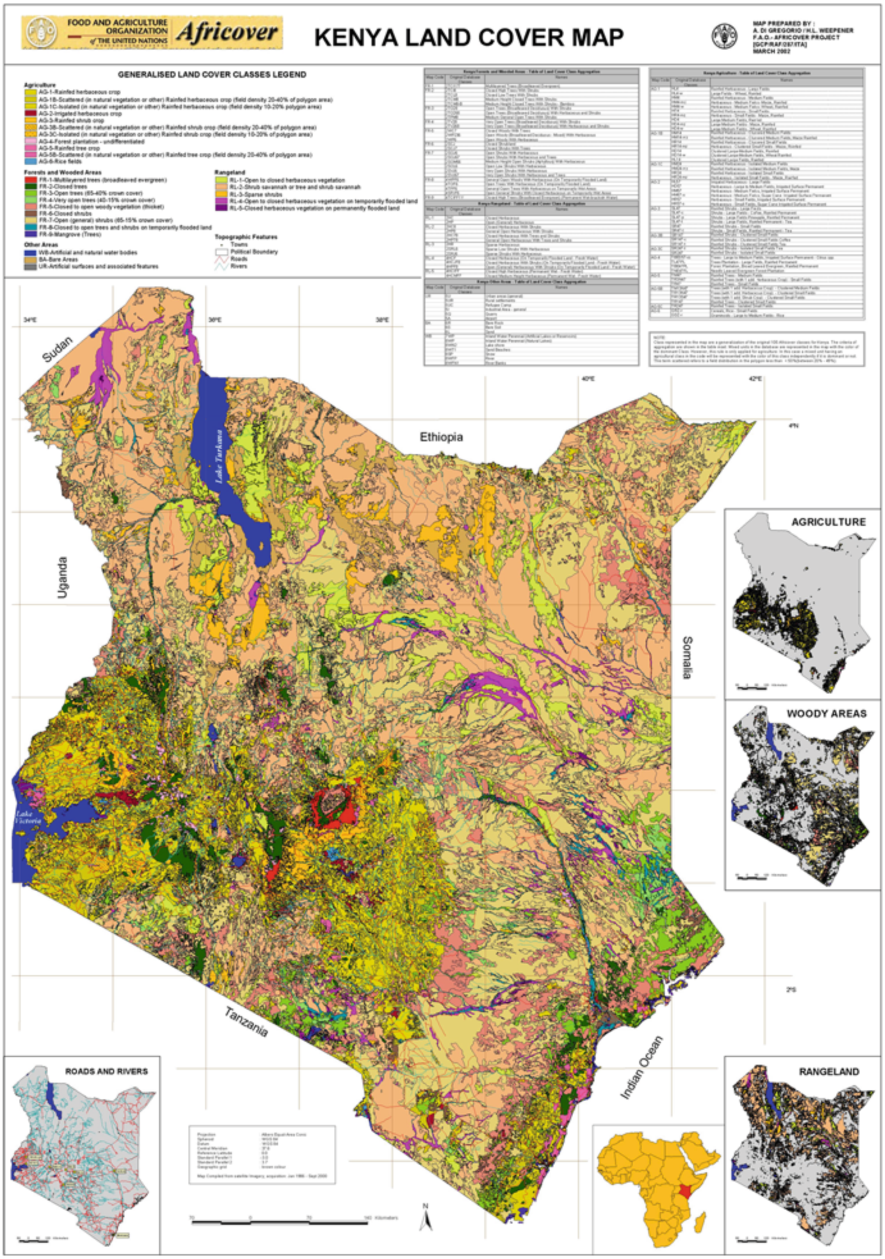


Plate 10.3 The F.A.O Africover Programme produced a digital georeferenced land cover database for 10 African countries (8.5 M km^2) at 1:200,000 scale (1:100,000 for small countries and specific areas), through the interpretation of Landsat images and applying the Land Cover Classification System (LCCS) methodology. The basic concepts are: (**a**) the ability to map very high level of details (tailored to the inherent characteristics of each country) maintaining at the same time a regional harmonization; (**b**) the data-base starts from local/national level to be later assembled at sub-regional/regional level. Aggregated land cover database for Kenya (15,000 polygons; 100 LC classes)

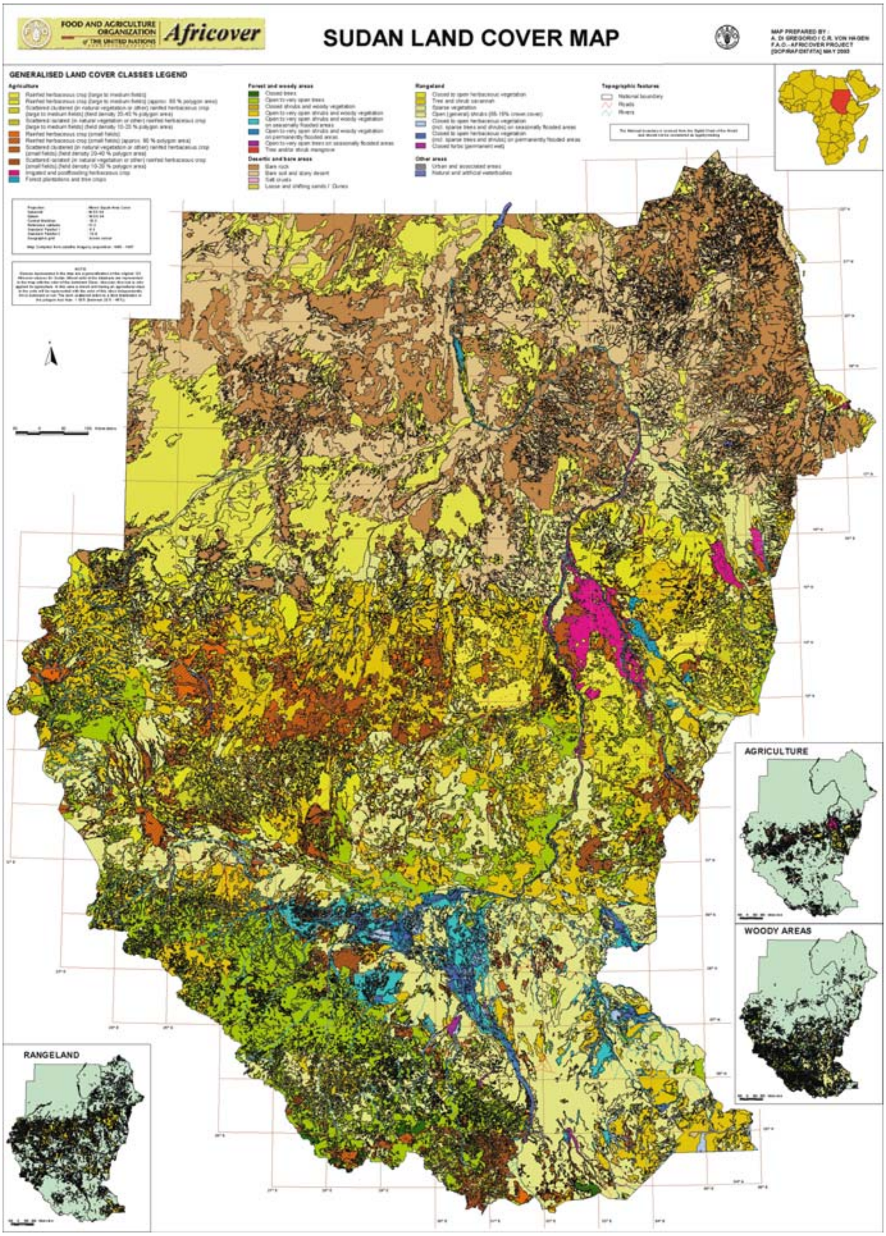


Plate 10.4 Aggregated land cover database for Sudan (30 000 polygons; 110 LC classes)

Index

Note: Locators with "f" represent figures and "t" represent tables

I

J

K

L

S